

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 22, 2003, 16:04:13 ; Search time 29.1819 Seconds
(without alignments)
1120.478 Million cell updates/sec

Title: US-09-887-855-2_COPY_22_227
Perfect score: 1115
Sequence: 1 ATGRLLSGQVCRGGTQRPC.....EEDAKKTFKESREAAALNLAY 206

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 1107863 seqs, 158726573 residues
Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq 19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES					
Result No.	Score	Query Match	Length DB ID	Description	
1	1115	100.0	374 21	AA93948	Amino acid sequenc
2	1115	100.0	374 22	AAE03651	Human extracellula
3	1115	100.0	374 23	ABG66680	Human novel polype
4	1115	100.0	374 23	ABB90203	Human polypeptide
5	1115	100.0	387 22	AAM25796	Human protein sequ
6	1105	99.1	374 21	AA91490	Human secreted pro
7	1101	98.7	382 20	AA913367	Amino acid sequenc
8	1101	98.7	382 22	AAU29033	Human PRO polypept
9	1101	98.7	382 22	AAB80235	Human PRO234 prote

10	1101	98.7	382	24	ABU69645	Novel human secret
11	1101	98.7	382	24	ABU71121	Human PRO234 prote
12	1101	98.7	382	24	ABU71468	Human PRO polypept
13	1101	98.7	382	24	ABU71914	Human secreted/tra
14	1101	98.7	382	24	ABU65578	Human secreted/tra
15	1101	98.7	382	24	ABU65911	Novel human secret
16	1101	98.7	382	24	ABU67368	Human secreted pro
17	1101	98.7	382	24	ABU67415	Human secreted/tra
18	1101	98.7	382	24	ABU64522	Human secreted/tra
19	1101	98.7	382	24	ABU65273	Human PRO polypept
20	1101	98.7	382	24	ABU58409	Human secreted/tra
21	1101	98.7	382	24	ABU55945	Human PRO polypept
22	1101	98.7	382	24	ABU56940	Human secreted/tra
23	1101	98.7	382	24	ABU54370	Human secreted/tra
24	1101	98.7	382	24	ABU10519	Human secreted/tra
25	561.5	50.4	273	21	AAB18913	A novel polypeptid
26	561.5	50.4	273	22	AAU12441	Human PRO1890 poly
27	561.5	50.4	273	22	AAB73309	Human C-type lecti
28	561.5	50.4	273	22	AAB87609	Human PRO1890. Ho
29	561.5	50.4	273	23	ABG95934	Human secreted/tra
30	561.5	50.4	273	23	ABB95554	Human angiogenesis
31	561.5	50.4	273	23	ABB84948	Human PRO1890 prot
32	561.5	50.4	273	24	ABU69084	Human PRO polypept
33	561.5	50.4	273	24	ABU69107	Human PRO polypept
34	561.5	50.4	273	24	ABU71589	Human secreted pol
35	561.5	50.4	273	24	ABU72035	Novel human secret
36	561.5	50.4	273	24	ABU72192	Human PRO polypept
37	561.5	50.4	273	24	ABU66839	Human PRO polypept
38	561.5	50.4	273	24	ABU67115	Human secreted/tra
39	561.5	50.4	273	24	ABU59920	Novel secreted and
40	534.5	47.9	274	23	ABP69211	Human polypeptide
41	504.5	45.2	232	22	AAB94192	Human protein sequ
42	275	24.7	102	22	ABG51278	Human liver peptid
43	275	24.7	102	22	ABB21788	Protein #3787 enco
44	275	24.7	102	22	AAM29941	Peptide #3978 enco
45	240	21.5	81	21	AA91643	Human secreted pro

ALIGNMENTS

RESULT 1
AA93948
ID AA93948 standard; Protein; 374 AA.
XX
AC AA93948;
XX
DT 03-OCT-2000 (first entry)
XX
DE Amino acid sequence of a lectin ss3939 polypeptide.
KW Human; lectin ss3939; chromosome 11; gene therapy.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..21
FT /note= "signal peptide"
FT Domain 22..227
FT /note= "extracellular coding region"
FT Domain 228..248
FT /note= "predicted transmembrane domain"
FT Domain 249..374
FT /note= "predicted cytoplasmic or intracellular domain"
XX
PN WO200039296-A1.
XX
PD 06-JUL-2000.
XX
PF 22-DEC-1999; 99WO-US30523.
XX
PR 23-DEC-1998; 98US-0113820.
XX

PA (IMMV) IMMUNEX CORP.
XX
PI Anderson DA;
XX
DR WPI; 2000-452394/39.
DR N-PSDB; AAA57382.
XX
PT ss3939 nucleic acids, polypeptides and antibodies, useful for
PT identifying human chromosome 11 and diseases associated with it -
XX
PS Claim 12; Page 8; 73pp; English.
XX
CC The present sequence represents a human lectin ss3939 polypeptide. The
CC polynucleotide sequence is a source of probes, which may be used
CC to identify nucleic acids encoding ss3939 proteins, to identify human
CC chromosome number 11, to map genes on human chromosome number 11, to
CC identify diseases associated with chromosome 11, as single-stranded
CC sense or antisense oligonucleotides to inhibit expression of
CC polypeptides encoded by the ss3939 gene, and for gene therapy. The
CC ss3939 polypeptides may be useful for developing treatments for
CC diseases (none specified) associated with defective or insufficient
CC amounts of the polypeptides. The antibodies may be useful for
CC detecting the presence of ss3939 polypeptides.
XX
SQ Sequence 374 AA;

Query Match 100.0%; Score 1115; DB 21; Length 374;
Best Local Similarity 100.0%; Pred. No. 3.5e-105;
Matches 206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGRLLSGQVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI ESEDEQKL 60
Db 22 ATGRLLSGQVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI ESEDEQKL 81
QY 61 IEKFIENLLPSDGDGFWIGLRRREKQSNSTACQDLYAWTDGSGISQFRNFWYVDEPSCGSEV 120
Db 82 IEKFIENLLPSDGDGFWIGLRRREKQSNSTACQDLYAWTDGSGISQFRNFWYVDEPSCGSEV 141
QY 121 CVVMYHQPSAPAGIGGYPYMFQWDDRCNMKNFNICKYSDEKPAVPSREAGEETELTPV 180
Db 142 CVVMYHQPSAPAGIGGYPYMFQWDDRCNMKNFNICKYSDEKPAVPSREAGEETELTPV 201
QY 181 LPEETQEDAKTKFESREAAALNLAY 206
Db 202 LPEETQEDAKTKFESREAAALNLAY 227

RESULT 2
AAE03651
ID AAE03651 standard; Protein; 374 AA.
XX
AC AAE03651;
XX
DT 06-AUG-2001 (first entry)
XX
DE Human extracellular matrix and cell adhesion molecule-15 (XMAD-15).
XX
KW Human; extracellular matrix and cell adhesion molecule; XMAD;
KW gene therapy; genetic disorder; adrenoleukodystrophy; leukaemia;
KW Down's syndrome; cystic fibrosis; Gaucher's disease; myotonic dystrophy;
KW sickle cell anaemia; thalassaemia; autoimmune disorder; adenocarcinoma;
KW inflammatory disorder; acquired immune deficiency syndrome; AIDS;
KW Addison's disease; allergy; anaemia; asthma; atherosclerosis; melanoma;
KW Crohn's disease; diabetes mellitus; atopic dermatitis; lymphoma; cancer;
KW glomerulonephritis; multiple sclerosis; Grave's disease; osteoarthritis;
KW osteoporosis; psoriasis; rheumatoid arthritis; ulcerative colitis;
KW infection; cell proliferative disorder; actinic keratosis; myeloma;
KW arteriosclerosis; neutropenic; anticonvulsant; antithyroid; nephrotropic;
KW neuroprotective; dermatological.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers

Peptide 1..24
Protein /label= Signal_peptide
Domain 25..374
Domain /note= "Mature human extracellular matrix and cell
Domain 46..63 adhesion molecule (XMAD)"
Domain /note= "C-type lectin domain"
Domain 163..176
Domain /note= "C-type lectin domain"
Domain 224..247
Domain /note= "Transmembrane motif"
Domain 328..348
Domain /note= "Transmembrane motif"
WO200142285-A2.
14-JUN-2001.
05-DEC-2000; 2000WO-US32990.
10-DEC-1999; 99US-0172852.
16-DEC-1999; 99US-0172354.
(INCY-) INCYTE GENOMICS INC.
Yue H, Tang YT, Lal P, Burford N, Azimzai Y, Patterson C;
Baughn MR, Lu DAM, Shah P, Au-Young J;
WPI; 2001-381632/40.
N-PSDB; AAD08059.
New human extracellular matrix and cell adhesion molecules and
polynucleotide sequences encoding them, useful for diagnosis,
prevention, treatment of genetic, autoimmune and cell proliferative
disorders -
Claim 1; Page 108-109; 135pp; English.
The present sequence is a human extracellular matrix and cell
adhesion molecule (XMAD). The XMAD is used for screening a compound for
effectiveness as an agonist or antagonist of XMAD. The identified agonist
or antagonist are used for treating a disease or condition associated
with decreased or increased expression of functional XMAD. The
polynucleotides encoding XMAD are useful in somatic or germline gene
therapy to correct a genetic deficiency, to express a conditionally
lethal gene product and to express a protein which affords protection
against intracellular parasites and also for diagnosis of disorders
associated with expression of XMAD. They are also used for generating
hybridisation probes useful in mapping the naturally occurring genomic
sequences (mice or rats) to model human diseases. Oligonucleotide or longer
fragments derived from the polynucleotide sequences may be used as
elements on a microarray. Antibodies which specifically bind XMAD may be
used for the diagnosis of disorders associated with the expression of
XMAD, or in assays to monitor patients being treated with XMAD. Diseases
diagnosed, prevented or treated include genetic disorders such as
adrenoleukodystrophy, Down's syndrome, cystic fibrosis, Gaucher's
disease, myotonic dystrophy, sickle cell anaemia, thalassaemia,
autoimmune/inflammatory disorders such as acquired immune deficiency
syndrome (AIDS), Addison's disease, diabetes mellitus, atopic asthma,
atherosclerosis, Crohn's disease, allergies, anaemia, asthma,
glomerulonephritis, multiple sclerosis, Grave's disease, osteoarthritis,
osteoporosis, psoriasis, rheumatoid arthritis, ulcerative colitis,
bacterial, fungal, parasitic, protozoal and helminthic infections and
cell proliferative disorders such as actinic keratosis, arteriosclerosis
and cancer including breast, bladder, bone marrow, brain and uterus
cancer, leukaemia, adenocarcinoma, lymphoma, melanoma and myeloma.
Sequence 374 AA;

Query Match 100.0%; Score 1115; DB 22; Length 374;
Best Local Similarity 100.0%; Pred. No. 3.5e-105;
Matches 206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEFAKEACRRDGGQLVSISEDEQKL 60
Db 22 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEFAKEACRRDGGQLVSISEDEQKL 81
QY 61 IEKFIENLLPSDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNMYVDEPSCGSEV 120
Db 82 IEKFIENLLPSDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNMYVDEPSCGSEV 141
QY 121 CVVMYHQPSAPAGIGGPFYMFQWDDRCNMKNFNICKYSDEKPAVPSREAEGETELTPV 180
Db 142 CVVMYHQPSAPAGIGGPFYMFQWDDRCNMKNFNICKYSDEKPAVPSREAEGETELTPV 201
QY 181 LPEETQEEADAKKTFKESREAAALNLAY 206
Db 202 LPEETQEEADAKKTFKESREAAALNLAY 227

RESULT 3
ABG66680
ID ABG66680 standard; Protein; 374 AA.
XX
AC ABG66680;
XX
DT 30-AUG-2002 (first entry)
XX
DE Human novel polypeptide #15.
XX

KW Human; inflammatory condition; shock; sepsis; immune response;
KW cancer; wound healing; central nervous system disease; haematopoiesis;
KW peripheral nervous system disease; amyotrophic lateral sclerosis; tendon;
KW myeloid cell disorder; lymphoid cell disorder; platelet disorder; bone;
KW cartilage; ligament; nerve tissue; ulcer; osteoporosis; osteoarthritis;
KW bone degenerative disorder; periodontal disease; reperfusion injury;
KW lung fibrosis; liver fibrosis; autoimmune disorder; bacterial infection;
KW allergic condition; thrombolysis; thrombosis; coagulation disorder;
KW fungal infection.

XX Homo sapiens.
OS
XX
PN WO200244340-A2.
XX
PD 06-JUN-2002.
XX
PF 30-NOV-2001; 2001WO-US47004.
XX
PR 30-NOV-2000; 2000US-0028952.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Goodrich RW, Liu C, Zhou P, Asundi V, Wang J, Wang D;
PI Yamazaki V, Ujwal ML, Drmanac RT;
XX
DR WPI; 2002-508509/54.
DR N-PSDB; ABK94904.
XX
PT Novel nucleic acids and polypeptides for diagnosis, treatment of
PT inflammatory, autoimmune, nervous system, myeloid or lymphoid cell
PT disorders, cancer and promoting wound healing -
XX
PS Claim 10; Page 579-580; 672pp; English.

XX The invention relates to human novel polynucleotides and associated
CC polypeptides. The polynucleotides and polypeptides are useful for
CC treating inflammatory conditions such as arthritis, nephritis, Crohn's
CC disease, ischaemia-reperfusion injury, shock, sepsis, immune responses
CC and cancer and for promoting wound healing. The sequences are used to
CC induce the proliferation of neural cells and regeneration of nerve and
CC brain tissue, and are useful for the treatment of central and peripheral
CC nervous system diseases and neuropathies, such as Alzheimer's disease,
CC Parkinson's disease, Huntington's disease and amyotrophic lateral
CC sclerosis. The sequences are involved in chemotactic or chemokinetic
CC activity, regulation of haematopoiesis, treatment of myeloid or lymphoid

CC cell disorders and platelet disorders such as thrombocytopenia,
CC regeneration of bone, cartilage, tendon, ligament and/or nerve tissue
CC growth, tissue repair, healing of burns, incisions, ulcers, treatment of
CC osteoporosis, osteoarthritis, bone degenerative disorders and periodontal
CC disease. The sequences of the invention are also useful for gut
CC protection or regeneration and treatment of lung or liver fibrosis,
CC reperfusion injury in various tissues, immune deficiencies and disorders
CC including severe combined immunodeficiency (SCID), bacterial or fungal
CC infections, autoimmune disorders e.g. multiple sclerosis and myasthenia
CC gravis, allergic conditions such as asthma, thrombolysis or thrombosis
CC and coagulation disorders. Sequences ABG66666-ABG66758 represent human
CC novel polypeptides of the invention.

XX
SQ Sequence 374 AA;
Query Match 100.0%; Score 1115; DB 23; Length 374;
Best Local Similarity 100.0%; Pred. NO. 3.5e-105;
Matches 206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEFAKEACRRDGGQLVSISEDEQKL 60
Db 22 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEFAKEACRRDGGQLVSISEDEQKL 81
QY 61 IEKFIENLLPSDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNMYVDEPSCGSEV 120
Db 82 IEKFIENLLPSDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNMYVDEPSCGSEV 141
QY 121 CVVMYHQPSAPAGIGGPFYMFQWDDRCNMKNFNICKYSDEKPAVPSREAEGETELTPV 180
Db 142 CVVMYHQPSAPAGIGGPFYMFQWDDRCNMKNFNICKYSDEKPAVPSREAEGETELTPV 201
QY 181 LPEETQEEADAKKTFKESREAAALNLAY 206
Db 202 LPEETQEEADAKKTFKESREAAALNLAY 227

RESULT 4

ABB90203
ID ABB90203 standard; Protein; 374 AA.
XX
AC ABB90203;
XX
DT 24-MAY-2002 (first entry)
XX
DE Human polypeptide SEQ ID NO 2579.
XX
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
KW vulnery; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein.
XX
OS Homo sapiens.
XX
PN WO200190304-A2.
XX
PD 29-NOV-2001.
XX
PF 18-MAY-2001; 2001WO-US16450.
XX
PR 19-MAY-2000; 2000US-205515P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Birse CE, Rosen CA;
XX
DR WPI; 2002-122018/16.
DR N-PSDB; ABL90612.
XX
PT Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
PT prevention of neural, immune system, muscular, reproductive,
PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
PT disorders -

XX PS Claim 11; SEQ ID NO 2579; 2081pp + Sequence Listing; English.

XX CC The invention relates to novel genes (ABL89449-ABL90853) and proteins (ABB89040-ABB90444) useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. The genes are isolated from a range of human tissues disclosed in the specification.

XX CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders such as myocardial ischaemias; (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f) infectious diseases such as viral, bacterial, fungal and parasitic infections.

XX CC Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 374 AA;

Query Match 100.0%; Score 1115; DB 23; Length 374;
Best Local Similarity 100.0%; Pred. No. 3.5e-105;
Matches 206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGRLLSGQPVCRGGTQPCYKVIYFHDTSRRLLNFEEAKEACRRDGGQLVSI ESEDEQKL 60
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy 61 IEKFIENLLPSDGDGFWIGLRRREEKQSNSTACQDLYAWTDGSI SQFRNWWYVDEPSCGSEV 120
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy 121 CVVMYHQPAPAGIGGYPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGETELTPV 180
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy 181 LPEETQEEADAKKTFKESREAAALNLAY 206
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 5
AAM25796
ID AAM25796 standard; Protein; 387 AA.
XX AC AAM25796;
XX 16-OCT-2001 (first entry)
XX Human protein sequence SEQ ID NO:1311.

KW Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
KW antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;
KW antibacterial; endocrine; cardiant; central nervous system; virucide;
KW anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;
KW antiaggregant; haemostatic; vulnery; antiulcer; osteopathic; eczema;
KW dermatological; antiallergic; antiasthmatic; antidiabetic; cytostatic;
KW neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
KW immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
KW antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
KW cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
KW genetic disease; haematopoietic disorder; platelet disorder; asthma;
KW thrombocytopaenia; osteoporosis; severe combined immunodeficiency;
KW allergic rhinitis; diabetes; multiple sclerosis; depression;
KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
XX neurological disorder.

OS Homo sapiens.
XX

PN WO200153455-A2.
XX 26-JUL-2001.
XX 22-DEC-2000; 2000WO-US35017.
XX 23-DEC-1999; 99US-0471275.
XX 21-JAN-2000; 2000US-0488725.
XX 25-APR-2000; 2000US-0552317.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Drmanac RT;
XX WPI; 2001-457603/49.
XX N-PSDB; AAH99737.

PT Isolated human polynucleotides encoding polypeptides, useful for the treatment and diagnosis of e.g. cancer, ulcers and HIV infection -

XX PS Claim 20; Page 272; 1217pp; English.

XX CC AAH99166 to AAH99904 encode the human proteins given in AAM25225 to AAM25963. The proteins can have activities based on the tissues and cells they are expressed in, such as: antiinflammatory; antirheumatic; antiarthritic; immunosuppressive; antibacterial; endocrine; cardiant; central nervous system; virucide; anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; antiaggregant; haemostatic; vulnery; antiulcer; osteopathic; dermatological; antiallergic; antiasthmatic; antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic; antiparkinsonian; and immunostimulant. The proteins and polynucleotides encoding them can be used in gene therapy, antisense therapy and vaccine production. The proteins and polynucleotides are useful for screening for agonists or antagonists of a protein and for the treatment and diagnosis of disorders associated with the activity of a protein e.g. inflammation, rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction, neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal infections, autoimmunity, genetic diseases, haematopoietic disorders, anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers, osteoporosis, severe combined immunodeficiency, eczema, allergic rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression, Alzheimer's disease, Parkinson's disease, neurodegenerative and neurological disorders.

XX SQ Sequence 387 AA;

Query Match 100.0%; Score 1115; DB 22; Length 387;
Best Local Similarity 100.0%; Pred. No. 3.6e-105;
Matches 206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGRLLSGQPVCRGGTQPCYKVIYFHDTSRRLLNFEEAKEACRRDGGQLVSI ESEDEQKL 60
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy 61 IEKFIENLLPSDGDGFWIGLRRREEKQSNSTACQDLYAWTDGSI SQFRNWWYVDEPSCGSEV 120
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy 121 CVVMYHQPAPAGIGGYPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGETELTPV 180
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy 181 LPEETQEEADAKKTFKESREAAALNLAY 206
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 6
AAY91490
ID AAY91490 standard; Protein; 374 AA.
XX AC AAY91490;
XX

DT 29-JUN-2000 (first entry)
XX Human secreted protein sequence encoded by gene 40 SEQ ID NO:163.
DE
XX
KW Human; secreted protein; diagnosis; cytostatic; immunosuppressive;
KW antiHIV; antiinflammatory; nootropic; neuroprotective; antiallergic;
KW osteopathic; antiarthritic; antibacterial; antidiabetic; antiasthma;
KW antipsoriatic; cardiant; gene therapy; cancer; neurological disorder;
KW immune disease; inflammation; blood disorder; tumour.
XX
OS Homo sapiens.
XX
XX WO200006698-A1.
PN
XX
XX 10-FEB-2000.
PD
XX
XX 29-JUL-1999; 99WO-US17130.
PF
XX 30-JUL-1998; 98US-0094657.
PR 05-AUG-1998; 98US-0095486.
PR 06-AUG-1998; 98US-0095454.
PR 06-AUG-1998; 98US-0095455.
PR 12-AUG-1998; 98US-0096319.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Komatsoulis GA, Rosen CA, Ruben SM, Duan R, Moore PA, Shi Y;
PI Lafleur D, Wei Y, Ni J, Florence KA, Young PE, Brewer LA;
PI Soppet DR, Endress GA, Ebner R, Olsen HS, Mucenski M;
XX
DR WPI; 2000-195282/17.
DR N-PSDB; AAA26385.
XX
PT New isolated human genes and the secreted polypeptides they encode,
PT useful for diagnosis and treatment of e.g. cancers, neurological
PT disorders, immune diseases, inflammation or blood disorders -
XX
PS Claim 11; Page 483-484; 634pp; English.
XX
CC The polynucleotide sequences given in AAA26346 to AAA26458 encode the
CC human secreted proteins given in AAY91451 to AAY91691. The human secreted
CC proteins can have activities based on the tissues and cells they are
CC expressed in. Examples of the activities are: cytostatic;
CC immunosuppressive; antiHIV; antiinflammatory; nootropic; neuroprotective;
CC antiallergic; osteopathic; antiarthritic; antibacterial; antidiabetic;
CC antiasthma; antipsoriatic; and cardiant. The polynucleotides and their
CC corresponding secreted proteins are useful for preventing, treating or
CC ameliorating medical conditions, e.g. by protein or gene therapy. Also
CC pathological conditions can be diagnosed by determining the amount of the
CC proteins in a sample or by determining the presence of mutations in the
CC polynucleotides. Specific uses are described for each of the
CC polynucleotides, based on which tissues they are most highly expressed
CC in, and include developing products for the diagnosis or treatment of
CC cancer, tumours, neurodegenerative disorders, developmental abnormalities
CC and foetal deficiencies, blood disorders, diseases of the immune system,
CC autoimmune diseases, hepatic and renal disease, inflammation,
CC allergies, Alzheimer's and behavioural disorders, schizophrenia,
CC osteoporosis, arthritis, infections, AIDS, spinal cord injuries,
CC transplant rejection, diabetes, asthma, sepsis, acne, psoriasis,
CC cardiovascular disorders, reproductive disorders, gastrointestinal
CC disorders, respiratory disorders and metabolic disorders. The proteins
CC or polynucleotides can also be used as food additives or preservatives.
CC The proteins are also useful for identifying their binding partners.
CC AAA26337 to AAA26345 and AAY91450 are sequences used in the
CC exemplification of the present invention.
XX
SQ Sequence 374 AA;

Query Match 99.1%; Score 1105; DB 21; Length 374;
Best Local Similarity 99.0%; Pred. No. 3.7e-104;
Matches 204; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNPFEEAKEACRRDGGQLVSI ESEDEQKL 60

Db 22 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNPFEEAKEACRRDGGQLVSI ESEDEQKL 81
QY 61 IEKFIENLLPSDGDGDFWIGLRRREEKQSNSTACODLYAWTDGSI SQFRNWWYVDEPSCGSEV 120
Db 82 IEXFIENLLPSDGDGDFWIGLRRREEKQSNSTXCQDLYAWTDGSI SQFRNWWYVDEPSCGSEV 141
QY 121 CVVMYHQPSPAPAGIGGPPYMFQWNDRCNMKNFNICKYSDEKPAVPSREAEGETELTTPV 180
Db 142 CVVMYHQPSPAPAGIGGPPYMFQWNDRCNMKNFNICKYSDEKPAVPSREAEGETELTTPV 201
QY 181 LPEETQEEDAKKTFKESREAAALNLAY 206
Db 202 LPEETQEEDAKKTFKESREAAALNLAY 227
RESULT 7
ID AAY13367
XX AAY13367 standard; Protein; 382 AA.
AC AAY13367;
XX
DT 25-JUN-1999 (first entry)
XX Amino acid sequence of protein PRO234.
DE Secreted protein; transmembrane protein; human; enterocolitis;
XX Zollinger-Ellison syndrome; gastrointestinal ulceration;
KW congenital microvillus atrophy; skin disease; cell growth;
KW abnormal keratinocyte differentiation; psoriasis; epithelial cancer;
KW Parkinson's disease; Alzheimer's disease; ALS; neuropathy;
KW fibromodulin; dermal scarring; Usher Syndrome; Atrophia areata;
KW anti-thrombotic; wound healing; tissue repair.
XX
OS Homo sapiens.
XX WO9914328-A2.
PN
XX
PD 25-MAR-1999.
XX
PF 16-SEP-1998; 98WO-US19330.
XX
XX 25-NOV-1997; 97US-0066840.
PR 17-SEP-1997; 97US-0059113.
PR 17-SEP-1997; 97US-0059115.
PR 17-SEP-1997; 97US-0059117.
PR 17-SEP-1997; 97US-0059119.
PR 17-SEP-1997; 97US-0059121.
PR 17-SEP-1997; 97US-0059122.
PR 17-SEP-1997; 97US-0059184.
PR 18-SEP-1997; 97US-0059263.
PR 18-SEP-1997; 97US-0059266.
PR 15-OCT-1997; 97US-0062125.
PR 17-OCT-1997; 97US-0062285.
PR 17-OCT-1997; 97US-0062287.
PR 21-OCT-1997; 97US-0063486.
PR 24-OCT-1997; 97US-0062814.
PR 24-OCT-1997; 97US-0062816.
PR 24-OCT-1997; 97US-0063045.
PR 24-OCT-1997; 97US-0063120.
PR 24-OCT-1997; 97US-0063121.
PR 24-OCT-1997; 97US-0063127.
PR 24-OCT-1997; 97US-0063128.
PR 27-OCT-1997; 97US-0063329.
PR 27-OCT-1997; 97US-0063327.
PR 28-OCT-1997; 97US-0063541.
PR 28-OCT-1997; 97US-0063542.
PR 28-OCT-1997; 97US-0063544.
PR 28-OCT-1997; 97US-0063549.
PR 28-OCT-1997; 97US-0063550.
PR 28-OCT-1997; 97US-0063564.
PR 29-OCT-1997; 97US-0063435.
PR 29-OCT-1997; 97US-0063704.

PR 29-OCT-1997; 97US-0063732.
PR 29-OCT-1997; 97US-0063738.
PR 29-OCT-1997; 97US-0063734.
PR 29-OCT-1997; 97US-0064215.
PR 29-OCT-1997; 97US-0063735.
PR 31-OCT-1997; 97US-0063870.
PR 31-OCT-1997; 97US-0064103.
PR 03-NOV-1997; 97US-0064248.
PR 07-NOV-1997; 97US-0064809.
PR 12-NOV-1997; 97US-0065186.
PR 17-NOV-1997; 97US-0065846.
PR 18-NOV-1997; 97US-0065693.
PR 21-NOV-1997; 97US-0066120.
PR 21-NOV-1997; 97US-0066364.
PR 24-NOV-1997; 97US-0066772.
PR 24-NOV-1997; 97US-0066466.
PR 24-NOV-1997; 97US-0066770.
PR 24-NOV-1997; 97US-0066511.
PR 24-NOV-1997; 97US-0066453.
XX
PA (GETH) GENENTECH INC.
XX
PI Chen J, Goddard A, Gurney AL, Pennica D, Wood WI, Yuan J;
XX
DR WPI; 1999-229533/19.
DR N-PSDB; AAX52238.
XX
PT New isolated human genes and polypeptides used in, e.g. treatment of
PT gastrointestinal ulceration
XX
PS Claim 12; Fig 50; 320pp; English.
XX
CC AAY13344-403 represent secreted and transmembrane human proteins.
CC The cDNA sequences are obtained from cDNA libraries, prepared from
CC fetal lung, fetal kidney, fetal brain, fetal liver and fetal retina.
CC The encoded polypeptides have specific uses based on their homology to
CC known polypeptides, e.g. PRO211 and PRO217 can be used for disorders
CC associated with the preservation and maintenance of gastrointestinal
CC mucosa and the repair of acute and chronic mucosal lesions
CC (e.g. enterocolitis, Zollinger-Ellison syndrome, gastrointestinal
CC ulceration and congenital microvillus atrophy), skin diseases associated
CC with abnormal keratinocyte differentiation (e.g. psoriasis, epithelial
CC cancers such as lung squamous cell carcinoma of the vulva and gliomas),
CC potent effects on cell growth and development, diseases related to
CC growth or survival of nerve cells including Parkinson's disease,
CC Alzheimer's disease, ALS, neuropathies or cancer. PRO265 can be used as
CC for fibromodulin, e.g. for reducing dermal scarring. PRO264 can be used
CC as a target for anti-tumor drugs. PRO533 may be used in the treatment
CC of Usher Syndrome or Atrophia areata; PRO269 can be used as an
CC anti-thrombotic agent; PRO287 polypeptides and portions may have
CC therapeutic applications in wound healing and tissue repair; PRO317 can
CC be used for treating problems of the kidney, uterus, endometrium, blood
CC vessels, or related tissue, e.g. in the heart of genital tract.
XX
SQ Sequence 382 AA;

Query Match 98.7%; Score 1101; DB 20; Length 382;
Best Local Similarity 96.3%; Pred. No. 9.6e-104;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 ATGRLLS-----GQVCRGGTQPCYKVIYFHDTSRRLNFEAKEACRRDGGQLVSI 52
Db |||||
22 ATGRLLSASDLDLRGGQPVCRGGTQPCYKVIYFHDTSRRLNFEAKEACRRDGGQLVSI 81

QY 53 ESEDEQKLEKFIENLLPSDGFNIGLRREEKQSNSTACQDLYAWTDGSIQFRNWWYVD 112
Db |||||
82 ESEDEQKLEKFIENLLPSDGFNIGLRREEKQSNSTACQDLYAWTDGSIQFRNWWYVD 141

QY 113 EPSCGSEVCVMYHQPSAPAGIGGPFYQWDDRCNMKNFICKYSDEKPAVPSREAEGE 172
Db |||||
142 EPSCGSEVCVMYHQPSAPAGIGGPFYQWDDRCNMKNFICKYSDEKPAVPSREAEGE 201

QY 173 ETELTTPVLPETQEEDAKKTFKESREAAALNLAY 206

Db 202 ETELTTPVLPETQEEDAKKTFKESREAAALNLAY 235
|||
RESULT 8
AAU29033
ID AAU29033 standard; Protein; 382 AA.
XX
AC AAU29033;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human PRO polypeptide sequence #10.
XX
KW PRO polypeptide; mammal; tumour; cancer; cattle; horse; sheep;
KW dog; cat; pig; goat; rabbit; tumour necrosis factor alpha; TNF-alpha;
KW blood; chondrocyte cell; cell proliferation; cell differentiation; colon;
KW adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200168848-A2.
XX
PD 20-SEP-2001.
XX
PF 28-FEB-2001; 2001WO-US06520.
XX
PR 01-MAR-2000; 2000WO-US05601.
PR 02-MAR-2000; 2000WO-US05841.
PR 03-MAR-2000; 2000US-187202P.
PR 06-MAR-2000; 2000US-186968P.
PR 14-MAR-2000; 2000US-189320P.
PR 14-MAR-2000; 2000US-189328P.
PR 15-MAR-2000; 2000WO-US06884.
PR 21-MAR-2000; 2000US-190828P.
PR 21-MAR-2000; 2000US-191007P.
PR 21-MAR-2000; 2000US-191048P.
PR 21-MAR-2000; 2000US-191314P.
PR 28-MAR-2000; 2000US-192655P.
PR 29-MAR-2000; 2000US-193032P.
PR 29-MAR-2000; 2000US-193053P.
PR 30-MAR-2000; 2000WO-US08439.
PR 04-APR-2000; 2000US-194449P.
PR 04-APR-2000; 2000US-194647P.
PR 11-APR-2000; 2000US-195975P.
PR 11-APR-2000; 2000US-196000P.
PR 11-APR-2000; 2000US-196187P.
PR 11-APR-2000; 2000US-196690P.
PR 11-APR-2000; 2000US-196820P.
PR 18-APR-2000; 2000US-198121P.
PR 18-APR-2000; 2000US-198585P.
PR 25-APR-2000; 2000US-199397P.
PR 25-APR-2000; 2000US-199550P.
PR 25-APR-2000; 2000US-199654P.
PR 03-MAY-2000; 2000US-201516P.
PR 17-MAY-2000; 2000WO-US13705.
PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.
PR 02-JUN-2000; 2000WO-US15264.
PR 05-JUN-2000; 2000US-209832P.
PR 28-JUL-2000; 2000WO-US20710.
PR 22-AUG-2000; 2000US-0644848.
PR 24-AUG-2000; 2000WO-US23328.
PR 08-NOV-2000; 2000WO-US30952.
PR 01-DEC-2000; 2000WO-US32678.
PR 20-DEC-2000; 2000WO-US34956.
XX
PA (GETH) GENENTECH INC.
XX
PI Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;
PI Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;
XX
DR WPI; 2001-602746/68.

DR N-PSDB; AAS45934.
XX Novel nucleic acids encoding PRO polypeptides, used to diagnose the
PT presence of tumours, such as prostate and breast tumours, in mammals and
PT to screen for modulators of the compounds -
XX
PS Claim 11; Fig 20; 774pp; English.
XX
CC Sequences AAU29024-AAU29328 represent PRO polypeptides of the invention.
CC The PRO polypeptides and their associated nucleic acids can be used to
CC detect the presence of a tumour in a mammal by comparing the level of
CC expression of a PRO polypeptide in a test sample of cells from the animal
CC and a control sample of normal cells, whereby a higher level of
CC expression in the test sample indicates the presence of a tumour in the
CC mammal. Mammals include dogs, cats, cattle, horses, sheep, pigs, goats
CC and rabbits but are preferably human. The polypeptides can be used to
CC stimulate tumour necrosis factor (TNF) alpha release from human blood,
CC when contacted with it. A specific polypeptide can be used to stimulate
CC the proliferation or differentiation of chondrocyte cells. The PRO
CC proteins can be used to determine the presence of tumours and also
CC susceptibility to tumour development, particularly adrenal, lung, colon,
CC breast, prostate, rectal, cervical, or liver tumours, in mammalian
CC subjects. The oligonucleotide probes specific for the PRO nucleic acids
CC can be used for genetic analysis of individuals with genetic disorders.
XX
SQ Sequence 382 AA;

Query Match 98.7%; Score 1101; DB 22; Length 382;
Best Local Similarity 96.3%; Pred. No. 9.6e-104;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
QY 1 ATGRLLS-----GQVCRGGTQPCYKVIYFHDTSRRLNFEAEACRRDGGQLVSI 52
Db 22 ATGRLLSASDLRLGGQPVCRGGTQPCYKVIYFHDTSRRLNFEAEACRRDGGQLVSI 81
QY 53 ESEDEQKLIKFIENLLPSDGFWDGFWIGLRRRREKQSNSTACQDLYAWTDGSIQFRNWWYVD 112
Db 82 ESEDEQKLIKFIENLLPSDGFWDGFWIGLRRRREKQSNSTACQDLYAWTDGSIQFRNWWYVD 141
QY 113 EPSCGSEVCVVMYHQPSAPAGIGGYPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGE 172
Db 142 EPSCGSEVCVVMYHQPSAPAGIGGYPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGE 201
QY 173 ETELTPVLPEETQEDAKKTFKESREAAALNLAY 206
Db 202 ETELTPVLPEETQEDAKKTFKESREAAALNLAY 235

RESULT 9
AAB80235
ID AAB80235 standard; Protein; 382 AA.
XX AAB80235;
AC
XX 24-APR-2001 (first entry)
DT
XX Human PRO234 protein.
DE
XX Human; PRO; dermatological; antipsoriatic; cytostatic; antiinflammatory;
KW antiparkinsonian nootropic; neuroprotective; vulnery; cardiant;
KW antiangiogenic; vasotropic; antiasthmatic; antirheumatic; cancer;
KW antiarthritic; antinfertility; antidiabetic; antiviral; diabetes;
KW ophthalmological; gene therapy; skin disease; gastrointestinal disorder;
KW ischaemia; inflammation.
XX
OS Homo sapiens.
XX
PN WO200104311-A1.
XX
PD 18-JAN-2001.
XX
PF 22-FEB-2000; 2000WO-US04414.
XX

PR 07-JUL-1999; 99US-0143048.
PR 26-JUL-1999; 99US-0145698.
PR 28-JUL-1999; 99US-0146222.
PR 08-SEP-1999; 99WO-US20594.
PR 13-SEP-1999; 99WO-US20944.
PR 15-SEP-1999; 99WO-US21090.
PR 15-SEP-1999; 99WO-US21547.
PR 05-OCT-1999; 99WO-US23089.
PR 29-NOV-1999; 99WO-US28214.
PR 30-NOV-1999; 99WO-US28313.
PR 16-DEC-1999; 99WO-US30095.
PR 20-DEC-1999; 99WO-US30911.
PR 20-DEC-1999; 99WO-US30999.
PR 05-JAN-2000; 99WO-US00219.
XX (GETH) GENENTECH INC.
PA Ashkenazi AJ, Botstein D, Desnoyers L, Eaton DL, Ferrara N;
XX Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME, Goddard A;
PI Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ, Kljavin IJ;
PI Mather JP, Pan J, Paoni NF, Roy MA, Stewart TA, Tumas D;
PI Williams PM, Wood WI;
XX WPI; 2001-081051/09.
DR N-PSDB; AAF72396.
XX
PT Sixty one nucleic acids encoding PRO polypeptides which are useful in
PT the treatment of skin diseases (e.g. psoriasis), cancers (e.g. lung
PT squamous cell carcinoma) and neurodegenerative diseases (e.g.
PT Alzheimer's disease) -
XX
PS Claim 1; Fig 50; 393pp; English.
XX
CC The present sequence is one of sixty one novel secreted and
CC transmembrane PRO polypeptides. The PRO polypeptides are
CC useful for treating skin diseases (e.g. psoriasis), cancers (e.g. lung
CC squamous cell carcinoma), gastrointestinal disorders (e.g.
CC enterocolitis), neurodegenerative diseases (e.g. Alzheimer's disease,
CC Parkinson's disease), wound repair, cardiovascular disorders (e.g.
CC endometrial bleeding angiogenesis, ischaemia such as coronary
CC ischaemia, atherosclerosis), inflammatory disorders (e.g. asthma,
CC rheumatoid arthritis, multiple sclerosis), infertility, AIDS and
CC diabetes and retinal disorders such as retinitis pigmentosum.
CC The PRO nucleic acids have applications in molecular biology, including
CC use as hybridization probes, and in chromosome and gene mapping.
XX
SQ Sequence 382 AA;

Query Match 98.7%; Score 1101; DB 22; Length 382;
Best Local Similarity 96.3%; Pred. No. 9.6e-104;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
QY 1 ATGRLLS-----GQVCRGGTQPCYKVIYFHDTSRRLNFEAEACRRDGGQLVSI 52
Db 22 ATGRLLSASDLRLGGQPVCRGGTQPCYKVIYFHDTSRRLNFEAEACRRDGGQLVSI 81
QY 53 ESEDEQKLIKFIENLLPSDGFWDGFWIGLRRRREKQSNSTACQDLYAWTDGSIQFRNWWYVD 112
Db 82 ESEDEQKLIKFIENLLPSDGFWDGFWIGLRRRREKQSNSTACQDLYAWTDGSIQFRNWWYVD 141
QY 113 EPSCGSEVCVVMYHQPSAPAGIGGYPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGE 172
Db 142 EPSCGSEVCVVMYHQPSAPAGIGGYPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGE 201
QY 173 ETELTPVLPEETQEDAKKTFKESREAAALNLAY 206
Db 202 ETELTPVLPEETQEDAKKTFKESREAAALNLAY 235

RESULT 10
ABU69645
ID ABU69645 standard; Protein; 382 AA.
XX

AC ABU69645;
XX
DT 05-JUN-2003 (first entry)
XX
DE Novel human secreted and transmembrane protein PRO234.
XX
KW Human; secreted and transmembrane protein; gene therapy; psoriasis;
KW enterocolitis; gastrointestinal ulceration; skin disease;
KW keratinocyte differentiation; epithelial cancer; Alzheimer's disease;
KW squamous cell carcinoma; Parkinson's disease; inflammatory disease;
KW amyotrophic lateral sclerosis; rheumatoid arthritis; asthma;
KW multiple sclerosis; organ failure; atherosclerosis; cardiac injury;
KW infertility; birth defect; premature aging; AIDS; cancer;
KW diabetic complication; wound repair; tissue re-growth.
XX
OS Homo sapiens.
XX
XX US2003017463-A1.
PN
XX
PD 23-JAN-2003.
XX
PF 11-JUL-2001; 2001US-0903640.
XX
PR 10-SEP-1998; 98WO-US18824.
PR 14-SEP-1998; 98WO-US19177.
PR 16-SEP-1998; 98WO-US19330.
PR 17-SEP-1998; 98WO-US19437.
PR 01-DEC-1998; 98WO-US25108.
PR 08-SEP-1999; 99WO-US20594.
PR 13-SEP-1999; 99WO-US20944.
PR 15-SEP-1999; 99WO-US21090.
PR 15-SEP-1999; 99WO-US21547.
PR 05-OCT-1999; 99WO-US23089.
PR 29-NOV-1999; 99WO-US28214.
PR 30-NOV-1999; 99WO-US28313.
PR 01-DEC-1999; 99WO-US28301.
PR 02-DEC-1999; 99WO-US28564.
PR 02-DEC-1999; 99WO-US28565.
PR 16-DEC-1999; 99WO-US30095.
PR 20-DEC-1999; 99WO-US30911.
PR 20-DEC-1999; 99WO-US30999.
PR 05-JAN-2000; 2000WO-US00219.
PR 11-FEB-2000; 2000WO-US03565.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US05004.
PR 02-MAR-2000; 2000WO-US05841.
PR 20-MAR-2000; 2000WO-US07377.
PR 30-MAR-2000; 2000WO-US08439.
PR 22-MAY-2000; 2000WO-US14042.
PR 02-JUN-2000; 2000WO-US15264.
PR 28-JUL-2000; 2000WO-US20710.
PR 24-AUG-2000; 2000WO-US23328.
PR 17-SEP-1997; 97US-059113P.
PR 17-SEP-1997; 97US-059115P.
PR 17-SEP-1997; 97US-059117P.
PR 17-SEP-1997; 97US-059119P.
PR 17-SEP-1997; 97US-059121P.
PR 17-SEP-1997; 97US-059122P.
PR 17-SEP-1997; 97US-059184P.
PR 18-SEP-1997; 97US-059263P.
PR 18-SEP-1997; 97US-059266P.
PR 15-OCT-1997; 97US-062125P.
PR 17-OCT-1997; 97US-062285P.
PR 17-OCT-1997; 97US-062287P.
PR 21-OCT-1997; 97US-063486P.
PR 24-OCT-1997; 97US-062814P.
PR 24-OCT-1997; 97US-062816P.
PR 24-OCT-1997; 97US-063045P.
PR 24-OCT-1997; 97US-063120P.
PR 24-OCT-1997; 97US-063121P.
PR 24-OCT-1997; 97US-063127P.
PR 24-OCT-1997; 97US-063128P.
PR 27-OCT-1997; 97US-063327P.

PR 27-OCT-1997; 97US-063329P.
PR 28-OCT-1997; 97US-063541P.
PR 28-OCT-1997; 97US-063542P.
PR 28-OCT-1997; 97US-063544P.
PR 28-OCT-1997; 97US-063549P.
PR 28-OCT-1997; 97US-063550P.
PR 28-OCT-1997; 97US-063564P.
PR 29-OCT-1997; 97US-063435P.
PR 29-OCT-1997; 97US-063704P.
PR 29-OCT-1997; 97US-063732P.
PR 29-OCT-1997; 97US-063734P.
PR 29-OCT-1997; 97US-063735P.
PR 29-OCT-1997; 97US-063738P.
PR 29-OCT-1997; 97US-064215P.
PR 31-OCT-1997; 97US-063870P.
PR 31-OCT-1997; 97US-064103P.
PR 03-NOV-1997; 97US-064248P.
PR 07-NOV-1997; 97US-064809P.
PR 12-NOV-1997; 97US-065186P.
PR 17-NOV-1997; 97US-065846P.
PR 18-NOV-1997; 97US-065693P.
PR 21-NOV-1997; 97US-066120P.
PR 21-NOV-1997; 97US-066364P.
PR 24-NOV-1997; 97US-066453P.
PR 24-NOV-1997; 97US-066466P.
PR 24-NOV-1997; 97US-066511P.
PR 24-NOV-1997; 97US-066770P.
PR 24-NOV-1997; 97US-066772P.
PR 25-NOV-1997; 97US-066840P.
PR 12-DEC-1997; 97US-069425P.
PR 04-JUN-1998; 98US-088026P.
PR 10-SEP-1998; 98US-099803P.
PR 14-SEP-1998; 98US-100262P.
PR 17-SEP-1998; 98US-100858P.
PR 13-OCT-1998; 98US-104080P.
PR 20-NOV-1998; 98US-109304P.
PR 22-DEC-1998; 98US-113296P.
PR 07-JUL-1999; 99US-143048P.
PR 26-JUL-1999; 99US-145698P.
PR 28-JUL-1999; 99US-146222P.
PR 18-SEP-2000; 2000US-0665350.
XX
PA (GETH) GENENTECH INC.
XX
PI Ashkenazi A, Botstein D, Desnoyers L, Eaton DL, Ferrara N;
PI Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME, Goddard A;
PI Godowski PJ, Grimaldi JC, Gurney AL, Hillan KJ, Kljavin IJ;
PI Mather JP, Pan J, Paoni NF, Roy MA, Stewart TA, Tumas D;
PI Williams PM, Wood WI;
XX
DR WPI; 2003-341586/32.
DR N-PSDB; ACA54920.
XX
PT New PRO polypeptides and nucleic acid molecules, useful in diagnosing
PT or treating inflammatory diseases, organ failure, atherosclerosis,
PT cardiac injury, infertility, cancer, AIDS, Alzheimer's disease or
PT Parkinson's disease -
XX
PS Claim 12; Fig 50; 473pp; English.
XX
CC The invention describes sixty one nucleic acids encoding PRO polypeptides
CC (secreted and transmembrane). The PRO polypeptides and nucleic acids are
CC useful in diagnosing or treating enterocolitis, gastrointestinal
CC ulceration, skin diseases associated with abnormal keratinocyte
CC differentiation, e.g. psoriasis or epithelial cancers such as squamous
CC cell carcinoma, Alzheimer's disease, Parkinson's disease, amyotrophic
CC lateral sclerosis, inflammatory diseases, e.g. rheumatoid arthritis,
CC asthma or multiple sclerosis, organ failure, atherosclerosis, cardiac
CC injury, infertility, birth defects, premature aging, AIDS, cancer,
CC diabetic complications, or mutations in general. The polypeptides are
CC also useful for wound repair and associated therapies concerned with
CC re-growth of tissue. The PRO polypeptides and nucleic acid molecules
CC are also useful in gene therapy, and as molecular weight markers for

CC protein electrophoresis purposes. The anti-PRO antibodies may be used
CC in diagnostic assays for PRO, or for the affinity purification of PRO
CC from recombinant cell culture or natural sources. This is the amino
CC acid sequence of a novel human PRO polypeptide.

SQ Sequence 382 AA;

Query Match 98.7%; Score 1101; DB 24; Length 382;
Best Local Similarity 96.3%; Pred. No. 9.6e-104;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 ATGRLLS-----GQVCRGGTQPCYKVIYFHDTSRRLNFEAEKACRRDGGQLVSI 52
Db |||||||

QY 53 ESEDEQKLIKFIENLLPSDGFWIGLRRREKQSNSTACQDLYAWTDGSIQFRNYYVD 112
Db |||||||

QY 113 EPSCGSEVCVMYHQPSAPAGIGGYPMFQWDDRCNMKNPFICKYSDEKPAVPSREAEGE 172
Db |||||||

QY 173 ETELTPVLPEETQEEADAKTKPKESREAAALNLAY 206
Db |||||||

QY 202 ETELTPVLPEETQEEADAKTKPKESREAAALNLAY 235

RESULT 11
ABU71121
ID ABU71121 standard; Protein; 382 AA.
XX
AC ABU71121;
XX
DT 10-JUN-2003 (first entry)
XX
DE Human PRO234 protein.
XX
KW Human; PRO; secreted; transmembrane; cytostatic; TNF-alpha; blood;
KW tumour necrosis factor alpha release; chondrocyte cell; proliferation;
KW differentiation; tumour; gene therapy.
XX
OS Homo sapiens.
XX
PN US2003036143-A1.
XX
PD 20-FEB-2003.
XX
PF 02-JUL-2002; 2002US-0187600.
XX
PR 16-SEP-1998; 98WO-US19330.
PR 07-OCT-1998; 98WO-US21141.
PR 01-DEC-1998; 98WO-US25108.
PR 08-MAR-1999; 99WO-US05028.
PR 14-MAY-1999; 99WO-US10733.
PR 02-JUN-1999; 99WO-US12252.
PR 01-SEP-1999; 99WO-US20111.
PR 15-SEP-1999; 99WO-US21090.
PR 01-DEC-1999; 99WO-US28301.
PR 02-DEC-1999; 99WO-US28551.
PR 30-DEC-1999; 99WO-US31274.
PR 05-JAN-2000; 2000WO-US00219.
PR 18-FEB-2000; 2000WO-US04341.
PR 18-FEB-2000; 2000WO-US04342.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US05004.
PR 01-MAR-2000; 2000WO-US05601.
PR 02-MAR-2000; 2000WO-US05841.
PR 15-MAR-2000; 2000WO-US06884.
PR 30-MAR-2000; 2000WO-US08439.
PR 17-MAY-2000; 2000WO-US13705.
PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.

PR 02-JUN-2000; 2000WO-US15264.
PR 28-JUL-2000; 2000WO-US20710.
PR 24-AUG-2000; 2000WO-US23328.
PR 08-NOV-2000; 2000WO-US30952.
PR 01-DEC-2000; 2000WO-US32678.
PR 20-DEC-2000; 2000WO-US34956.
PR 28-FEB-2001; 2001WO-US06520.
PR 01-JUN-2001; 2001WO-US17800.
PR 20-JUN-2001; 2001WO-US19692.
PR 29-JUN-2001; 2001WO-US21066.
PR 09-JUL-2001; 2001WO-US21735.
PR 29-AUG-2001; 2001WO-US27099.
PR 18-SEP-1997; 97US-059263P.
PR 18-SEP-1997; 97US-059266P.
PR 17-OCT-1997; 97US-062250P.
PR 21-OCT-1997; 97US-063486P.
PR 24-OCT-1997; 97US-063120P.
PR 24-OCT-1997; 97US-063121P.
PR 28-OCT-1997; 97US-063540P.
PR 28-OCT-1997; 97US-063541P.
PR 28-OCT-1997; 97US-063544P.
PR 28-OCT-1997; 97US-063564P.
PR 29-OCT-1997; 97US-063734P.
PR 31-OCT-1997; 97US-063870P.
PR 31-OCT-1997; 97US-064103P.
PR 13-NOV-1997; 97US-065311P.
PR 21-NOV-1997; 97US-066120P.
PR 24-NOV-1997; 97US-066466P.
PR 24-NOV-1997; 97US-066772P.
PR 11-DEC-1997; 97US-069335P.
PR 12-DEC-1997; 97US-069425P.
PR 17-DEC-1997; 97US-069870P.
PR 18-DEC-1997; 97US-068017P.
PR 10-MAR-1998; 98US-077450P.
PR 11-MAR-1998; 98US-077649P.
PR 20-MAR-1998; 98US-078886P.
PR 20-MAR-1998; 98US-078939P.
PR 27-MAR-1998; 98US-079664P.
PR 27-MAR-1998; 98US-079786P.
PR 31-MAR-1998; 98US-080107P.
PR 31-MAR-1998; 98US-080194P.
PR 01-APR-1998; 98US-080327P.
PR 01-APR-1998; 98US-080333P.
PR 08-APR-1998; 98US-081049P.
PR 08-APR-1998; 98US-081070P.
PR 09-APR-1998; 98US-081195P.
PR 15-APR-1998; 98US-081838P.
PR 21-APR-1998; 98US-082568P.
PR 21-APR-1998; 98US-082569P.
PR 22-APR-1998; 98US-082704P.
PR 22-APR-1998; 98US-082797P.
PR 28-APR-1998; 98US-083322P.
PR 29-APR-1998; 98US-083495P.
PR 29-APR-1998; 98US-083496P.
PR 29-APR-1998; 98US-083499P.
PR 29-APR-1998; 98US-083559P.
PR 05-MAY-1998; 98US-084366P.
PR 06-MAY-1998; 98US-084414P.
PR 07-MAY-1998; 98US-084639P.
PR 07-MAY-1998; 98US-084640P.
PR 07-MAY-1998; 98US-084643P.
PR 15-MAY-1998; 98US-085579P.
PR 15-MAY-1998; 98US-085580P.
PR 15-MAY-1998; 98US-085582P.
PR 15-MAY-1998; 98US-085700P.
PR 18-MAY-1998; 98US-086023P.
PR 22-MAY-1998; 98US-086392P.
PR 22-MAY-1998; 98US-086486P.
PR 28-MAY-1998; 98US-087098P.
PR 28-MAY-1998; 98US-087208P.
PR 02-JUN-1998; 98US-087609P.
PR 02-JUN-1998; 98US-087759P.

PR 03-JUN-1998; 98US-087827P.
PR 04-JUN-1998; 98US-088025P.
PR 04-JUN-1998; 98US-088028P.
PR 04-JUN-1998; 98US-088029P.
PR 04-JUN-1998; 98US-088033P.
PR 04-JUN-1998; 98US-088326P.
PR 05-JUN-1998; 98US-088167P.
PR 05-JUN-1998; 98US-088202P.
PR 05-JUN-1998; 98US-088212P.
PR 05-JUN-1998; 98US-088217P.
PR 09-JUN-1998; 98US-088655P.
PR 10-JUN-1998; 98US-088722P.
PR 10-JUN-1998; 98US-088738P.
PR 10-JUN-1998; 98US-088740P.
PR 10-JUN-1998; 98US-088811P.
PR 10-JUN-1998; 98US-088824P.
PR 10-JUN-1998; 98US-088825P.
PR 10-JUN-1998; 98US-088826P.
PR 11-JUN-1998; 98US-088861P.
PR 11-JUN-1998; 98US-088863P.
PR 11-JUN-1998; 98US-088876P.
PR 12-JUN-1998; 98US-089090P.
PR 12-JUN-1998; 98US-089105P.
PR 16-JUN-1998; 98US-089512P.
PR 16-JUN-1998; 98US-089514P.
PR 17-JUN-1998; 98US-089538P.
PR 17-JUN-1998; 98US-089598P.
PR 17-JUN-1998; 98US-089653P.
PR 18-JUN-1998; 98US-089908P.
PR 19-JUN-1998; 98US-089952P.
PR 22-JUN-1998; 98US-090246P.
PR 22-JUN-1998; 98US-090252P.
PR 22-JUN-1998; 98US-090254P.
PR 24-JUN-1998; 98US-090429P.
PR 24-JUN-1998; 98US-090435P.
PR 24-JUN-1998; 98US-090444P.
PR 24-JUN-1998; 98US-090461P.
PR 24-JUN-1998; 98US-090535P.
PR 24-JUN-1998; 98US-090540P.
PR 25-JUN-1998; 98US-090676P.
PR 25-JUN-1998; 98US-090678P.
PR 25-JUN-1998; 98US-090688P.
PR 25-JUN-1998; 98US-090690P.
PR 25-JUN-1998; 98US-090694P.
PR 25-JUN-1998; 98US-090695P.
PR 25-JUN-1998; 98US-090696P.
PR 26-JUN-1998; 98US-090862P.
PR 26-JUN-1998; 98US-090863P.
PR 26-JUN-1998; 98US-091010P.
PR 01-JUL-1998; 98US-091359P.
PR 01-JUL-1998; 98US-091544P.
PR 02-JUL-1998; 98US-091478P.
PR 02-JUL-1998; 98US-091486P.
PR 02-JUL-1998; 98US-091626P.
PR 02-JUL-1998; 98US-091628P.
PR 02-JUL-1998; 98US-091632P.
PR 24-JUL-1998; 98US-094006P.
PR 04-AUG-1998; 98US-095282P.
PR 10-AUG-1998; 98US-095998P.
PR 10-AUG-1998; 98US-096012P.
PR 17-AUG-1998; 98US-096757P.
PR 17-AUG-1998; 98US-096766P.
PR 17-AUG-1998; 98US-096867P.
PR 17-AUG-1998; 98US-096891P.
PR 17-AUG-1998; 98US-096897P.
PR 18-AUG-1998; 98US-096949P.
PR 18-AUG-1998; 98US-096959P.
PR 18-AUG-1998; 98US-097022P.
PR 26-AUG-1998; 98US-097952P.
PR 26-AUG-1998; 98US-097954P.
PR 26-AUG-1998; 98US-097955P.
PR 26-AUG-1998; 98US-097971P.
PR 26-AUG-1998; 98US-097974P.

PR 26-AUG-1998; 98US-098014P.
PR 01-SEP-1998; 98US-098716P.
PR 01-SEP-1998; 98US-098723P.
PR 02-SEP-1998; 98US-098803P.
PR 02-SEP-1998; 98US-098821P.
PR 02-SEP-1998; 98US-098843P.
PR 09-SEP-1998; 98US-099602P.
PR 10-SEP-1998; 98US-099741P.
PR 10-SEP-1998; 98US-099754P.
PR 10-SEP-1998; 98US-099763P.
PR 10-SEP-1998; 98US-099812P.

Query Match 98.7%; Score 1101; DB 24; Length 382;
Best Local Similarity 96.3%; Pred.No. 9.6e-104;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 ATGRLLS-----GQVRCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
|||
Db 22 ATGRLLSASDLRLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81
|||

QY 53 ESEDEQKLIKFIENLLPSDGDFTWIGLRREEKQSNSTACQDLVYAWTDGSI SQFRNYYVD 112
|||
Db 82 ESEDEQKLIKFIENLLPSDGDFTWIGLRREEKQSNSTACQDLVYAWTDGSI SQFRNYYVD 141
|||

QY 113 EPSCGSEVCVVMYHQPSAPAGIGGPPYMFQWDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
|||
Db 142 EPSCGSEVCVVMYHQPSAPAGIGGPPYMFQWDDRCNMKNNFICKYSDEKPAVPSREAEGE 201
|||

QY 173 ETELTPVLPETQEEADAKTKFKESREAAALNLAY 206
|||
Db 202 ETELTPVLPETQEEADAKTKFKESREAAALNLAY 235
|||

RESULT 12

ABU71468
ID ABU71468 standard; Protein; 382 AA.
XX
AC ABU71468;
XX
DT 10-JUN-2003 (first entry)
XX
DE Human PRO polypeptide #24.
XX
KW Human; secreted and transmembrane protein; PRO polypeptide; cancer;
KW Alzheimer's disease; ischaemia; cytostatic; neurotropic; vasotropic;
KW neuroprotective.
XX
OS Homo sapiens.
XX
PN US2002192659-A1.
XX
PD 19-DEC-2002.
XX
PF 10-JUL-2001; 2001US-0902853.
XX
PR 10-SEP-1998; 98WO-US18824.
PR 14-SEP-1998; 98WO-US19177.
PR 16-SEP-1998; 98WO-US19330.
PR 17-SEP-1998; 98WO-US19437.
PR 01-DEC-1998; 98WO-US25108.
PR 08-SEP-1999; 99WO-US20594.
PR 13-SEP-1999; 99WO-US20944.
PR 15-SEP-1999; 99WO-US21090.
PR 15-SEP-1999; 99WO-US21547.
PR 05-OCT-1999; 99WO-US23089.
PR 01-DEC-1999; 99WO-US28301.
PR 02-DEC-1999; 99WO-US28564.
PR 02-DEC-1999; 99WO-US28565.
PR 16-DEC-1999; 99WO-US30095.
PR 20-DEC-1999; 99WO-US30911.
PR 20-DEC-1999; 99WO-US30999.
PR 05-JAN-2000; 2000WO-US00219.
PR 11-FEB-2000; 2000WO-US03565.

PR 22-FEB-2000; 2000WO-US04414.
PR 28-JUL-2000; 2000WO-US20710.
PR 24-AUG-2000; 2000WO-US23328.
PR 17-SEP-1997; 97US-059113P.
PR 17-SEP-1997; 97US-059115P.
PR 17-SEP-1997; 97US-059117P.
PR 18-SEP-1997; 97US-059266P.
PR 15-OCT-1997; 97US-062125P.
PR 17-OCT-1997; 97US-062285P.
PR 17-OCT-1997; 97US-062287P.
PR 21-OCT-1997; 97US-063486P.
PR 24-OCT-1997; 97US-062814P.
PR 24-OCT-1997; 97US-062816P.
XX
PA (GETH) GENENTECH INC.
XX

PI Ashkenazi A, Botstein D, Desnoyers L, Eaton DL, Ferrara N;

PI Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME, Goddard A;
PI Godowski PJ, Grimaldi JC, Gurney AL, Hillan KJ, Kljavin IJ;
PI Mather JP, Pan J, Paoni NF, Roy MA, Stewart TA, Tumas D;
PI Williams PM, Wood WI;
XX

DR WPI; 2003-361832/34.
DR N-PSDB; ACA58405.

XX
PT New isolated nucleic acid encoding a PRO polypeptide, e.g. PRO245 or
PT PRO1868, useful in molecular biology, chromosome and gene mapping, in
PT generating antisense RNA and DNA, and in gene therapy -

XX
PS Claim 12; Fig 50; 474pp; English.

XX
CC The present invention relates to the isolation of novel human secreted
CC and transmembrane proteins (PRO polypeptides), and the polynucleotide
CC sequences encoding them. The polynucleotide sequences are useful in
CC molecular biology, as hybridisation probes, in chromosome and gene
CC mapping, in generating antisense RNA and DNA, and in gene therapy. The
CC polynucleotide sequences may also be used in preparing PRO polypeptides
CC by recombinant techniques, and in generating either transgenic animals
CC or knock-out animals which, in turn, are useful in the development and
CC screening of therapeutically useful reagents. The PRO polypeptides or
CC their antibodies are useful in preparing a medicament for treating a
CC condition responsive to the polypeptide or antibody, such as cancer,
CC Alzheimer's disease or ischaemia, and in various diagnostic assays.
CC ABU71445-ABU71505 represent human PRO polypeptides of the invention.

XX
SQ Sequence 382 AA;

Query Match 98.7%; Score 1101; DB 24; Length 382;
Best Local Similarity 96.3%; Pred. No. 9.6e-104;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 ATGRLLS-----GQVCRGGTQPCYKVIYFHTSRRLNFEAEKACRRDGGQLVSI 52
Db |||||
22 ATGRLLSASDLRLRGQPVCRGGTQPCYKVIYFHTSRRLNFEAEKACRRDGGQLVSI 81
QY 53 ESEDEQKLIKFIENLLPSDGFWIGLRRRREKQSNSTACQDLYAWTDGSIQFRNWIYVD 112
Db |||||
82 ESEDEQKLIKFIENLLPSDGFWIGLRRRREKQSNSTACQDLYAWTDGSIQFRNWIYVD 141
QY 113 EPSCGSEVCVMYTHQPSAPAGIGGYPYMFQWDDRCNMKNFICKYSDEKPAVPSREAAGE 172
Db |||||
142 EPSCGSEVCVMYTHQPSAPAGIGGYPYMFQWDDRCNMKNFICKYSDEKPAVPSREAAGE 201
QY 173 ETELTTPVLPETQEEDAKKTFKESREAAALNLAY 206
Db |||||
202 ETELTTPVLPETQEEDAKKTFKESREAAALNLAY 235

RESULT 13
ABU71914
ID ABU71914 standard; Protein; 382 AA.
XX
AC ABU71914;

XX
DT 12-JUN-2003 (first entry)
XX
DE Human secreted/transmembrane protein PRO2334.
XX
KW Human; secreted protein; transmembrane protein; PRO;
KW gene therapy; chromosome identification; chromosome marker.
XX
OS Homo sapiens.
XX
PN US2003003530-A1.
XX
PD 02-JAN-2003.
XX
PF 11-JUL-2001; 2001US-0904011.
XX
PR 10-SEP-1998; 98WO-US18824.
PR 14-SEP-1998; 98WO-US19177.
PR 16-SEP-1998; 98WO-US19330.
PR 17-SEP-1998; 98WO-US19437.
PR 01-DEC-1998; 98WO-US25108.
PR 08-SEP-1999; 99WO-US20594.
PR 13-SEP-1999; 99WO-US20944.
PR 15-SEP-1999; 99WO-US21090.
PR 15-SEP-1999; 99WO-US21547.
PR 05-OCT-1999; 99WO-US23089.
PR 29-NOV-1999; 99WO-US28214.
PR 30-NOV-1999; 99WO-US28313.
PR 01-DEC-1999; 99WO-US28301.
PR 02-DEC-1999; 99WO-US28564.
PR 02-DEC-1999; 99WO-US28565.
PR 16-DEC-1999; 99WO-US30095.
PR 20-DEC-1999; 99WO-US30911.
PR 20-DEC-1999; 99WO-US30999.
PR 05-JAN-2000; 2000WO-US00219.
PR 11-FEB-2000; 2000WO-US03565.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US05004.
PR 02-MAR-2000; 2000WO-US05841.
PR 20-MAR-2000; 2000WO-US07377.
PR 30-MAR-2000; 2000WO-US08439.
PR 22-MAY-2000; 2000WO-US14042.
PR 02-JUN-2000; 2000WO-US15264.
PR 28-JUL-2000; 2000WO-US20710.
PR 24-AUG-2000; 2000WO-US23328.
PR 17-SEP-1997; 97US-059113P.
PR 17-SEP-1997; 97US-059115P.
PR 17-SEP-1997; 97US-059117P.
PR 17-SEP-1997; 97US-059119P.
PR 17-SEP-1997; 97US-059121P.
PR 17-SEP-1997; 97US-059122P.
PR 17-SEP-1997; 97US-059184P.
PR 18-SEP-1997; 97US-059263P.
PR 18-SEP-1997; 97US-059266P.
PR 15-OCT-1997; 97US-062125P.
PR 17-OCT-1997; 97US-062285P.
PR 17-OCT-1997; 97US-062287P.
PR 21-OCT-1997; 97US-063486P.
PR 24-OCT-1997; 97US-062814P.
PR 24-OCT-1997; 97US-062816P.
PR 24-OCT-1997; 97US-063045P.
PR 24-OCT-1997; 97US-063120P.
PR 24-OCT-1997; 97US-063121P.
PR 24-OCT-1997; 97US-063127P.
PR 24-OCT-1997; 97US-063128P.
PR 27-OCT-1997; 97US-063327P.
PR 27-OCT-1997; 97US-063329P.
PR 28-OCT-1997; 97US-063541P.
PR 28-OCT-1997; 97US-063542P.
PR 28-OCT-1997; 97US-063544P.
PR 28-OCT-1997; 97US-063549P.
PR 28-OCT-1997; 97US-063550P.
PR 28-OCT-1997; 97US-063564P.

PR 29-OCT-1997; 97US-063435P.
PR 29-OCT-1997; 97US-063704P.
PR 29-OCT-1997; 97US-063732P.
PR 29-OCT-1997; 97US-063734P.
PR 29-OCT-1997; 97US-063735P.
PR 29-OCT-1997; 97US-063738P.
PR 29-OCT-1997; 97US-064215P.
PR 31-OCT-1997; 97US-063870P.
PR 31-OCT-1997; 97US-064103P.
PR 03-NOV-1997; 97US-064248P.
PR 07-NOV-1997; 97US-064809P.
PR 12-NOV-1997; 97US-065186P.
PR 17-NOV-1997; 97US-065846P.
PR 18-NOV-1997; 97US-065893P.
PR 21-NOV-1997; 97US-066120P.
PR 21-NOV-1997; 97US-066364P.
PR 24-NOV-1997; 97US-066453P.
PR 24-NOV-1997; 97US-066466P.
PR 24-NOV-1997; 97US-066511P.
PR 24-NOV-1997; 97US-066770P.
PR 24-NOV-1997; 97US-066772P.
PR 18-SEP-2000; 2000US-0665350.
XX
PA (GETH) GENENTECH INC.
XX
PI Ashkenazi A, Botstein D, Desnoyers L, Eaton DL, Ferrara N;
PI Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME, Goddard A;
PI Godowski PJ, Grimaldi JC, Gurney AL, Hillan KJ, Kljavin IJ;
PI Mather JP, Pan J, Paoni NF, Roy MA, Stewart TA, Tumas D;
PI Williams PM, Wood WI;
XX
DR WPI; 2003-329602/31.
DR N-PSDB; ACA60112.
XX
PT New transmembrane polypeptides and nucleic acids encoding the
PT polypeptides, useful in gene therapy, in chromosome identification, as
PT chromosome markers, in generating probes and in tissue typing -
XX
PS Claim 12; Fig 50; 484pp; English.
XX
CC The invention relates to an isolated nucleic acid with at least 80%
CC nucleic acid sequence identity to a nucleotide sequence encoding one of
CC 61 secreted/transmembrane polypeptides, or PRO polypeptides or encoding a
CC PRO protein extracellular domain. Also included are a vector comprising
CC the PRO nucleic acid, a host cell comprising the vector, producing a PRO
CC polypeptide (by culturing the host cell for the expression of the PRO
CC polypeptide, and recovering the PRO polypeptide from the cell culture),
CC an isolated PRO polypeptide (having at least 80% sequence identity
CC to: (a) an amino acid sequence selected from the 61 PRO proteins;
CC (b) an amino acid sequence encoded by a nucleic acid molecule deposited
CC with an ATCC number (detailed in the specification); or (c) an
CC extracellular domain of a PRO polypeptide or to a PRO polypeptide lacking
CC its associated signal peptide), a chimaeeric molecule comprising a PRO
CC polypeptide of fused to a heterologous amino acid sequence, an anti-PRO
CC antibody, detecting a PRO245 or PRO1868 in a sample suspected of
CC containing the polypeptide, linking a bioactive molecule to a cell
CC expressing a PRO245 or PRO1868 and modulating at least one biological
CC activity of a cell expressing a PRO245 or PRO1868. Nucleic acids which
CC encode PRO can be used to generate either transgenic animals or knock-out
CC animals which may be used in the development and screening of
CC therapeutically useful reagents. The nucleic acids may also be used in
CC gene therapy, in chromosome identification, as chromosome markers, or in
CC generating probes. The PRO polypeptides are useful as molecular markers
CC for protein electrophoresis, and the isolated nucleic acids may be used
CC for recombinantly expressing those markers. The PRO polypeptides and
CC nucleic acids may also be used in tissue typing. Anti-PRO antibodies
CC are useful in diagnostic assays for PRO, and in affinity purification
CC of PRO from recombinant cell culture or natural sources. The
XX present sequence represents a PRO protein.
SQ Sequence 382 AA;
Query Match 98.7%; Score 1101; DB 24; Length 382;

Best Local Similarity 96.3%; Pred. No. 9.6e-104;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
Qy 1 ATGRLLS-----GQPVCRCGGTQRPCYKVYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
Db |||||
22 ATGRLLSASDLRLGGQPVCRCGGTQRPCYKVYFHDTSRRLNFEEAKEACRRDGGQLVSI 81
Qy 53 ESEDEQKLIKFIENLLPSDGFWDGFWIGLRRREKQSNSTACQDLYAWTDGSIQFRNMYVD 112
Db |||||
82 ESEDEQKLIKFIENLLPSDGFWDGFWIGLRRREKQSNSTACQDLYAWTDGSIQFRNMYVD 141
Qy 113 EPSCGSEVCVVMYHQPSAPAGIGGYPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGE 172
Db |||||
142 EPSCGSEVCVVMYHQPSAPAGIGGYPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGE 201
Qy 173 ETELTTPVLPEETQEEDAKKTFKESREAAALNLAY 206
Db |||||
202 ETELTTPVLPEETQEEDAKKTFKESREAAALNLAY 235
RESULT 14
ABU65578
ID ABU65578 standard; Protein; 382 AA.
XX
AC ABU65578;
XX
DT 19-MAY-2003 (first entry)
XX
DE Human secreted/transmembrane protein, SEQ ID 20.
XX
KW Human; PRO; secreted protein; transmembrane protein;
KW cytotstatic; antiarthritic; osteopathic; adrenal tumour; lung tumour;
KW colon tumour; breast tumour; prostate tumour; rectal tumour;
KW cervical tumour; liver tumour; TNF-alpha release; arthritis;
KW tumour necrosis factor alpha; chondrocyte cell; bone disorder;
KW cartilage disorder; sports injury.
XX
OS Homo sapiens.
XX
PN US2003036156-A1.
XX
PD 20-FEB-2003.
XX
PF 02-JUL-2002; 2002US-0188767.
XX
PR 16-SEP-1998; 98WO-US19330.
PR 07-OCT-1998; 98WO-US21141.
PR 01-DEC-1998; 98WO-US25108.
PR 08-MAR-1999; 99WO-US05028.
PR 14-MAY-1999; 99WO-US10733.
PR 02-JUN-1999; 99WO-US12252.
PR 01-SEP-1999; 99WO-US20111.
PR 15-SEP-1999; 99WO-US21090.
PR 01-DEC-1999; 99WO-US28301.
PR 02-DEC-1999; 99WO-US28551.
PR 30-DEC-1999; 99WO-US31274.
PR 05-JAN-2000; 2000WO-US00219.
PR 18-FEB-2000; 2000WO-US04341.
PR 18-FEB-2000; 2000WO-US04342.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US05004.
PR 01-MAR-2000; 2000WO-US05601.
PR 02-MAR-2000; 2000WO-US05841.
PR 15-MAR-2000; 2000WO-US06884.
PR 30-MAR-2000; 2000WO-US08439.
PR 17-MAY-2000; 2000WO-US13705.
PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.
PR 02-JUN-2000; 2000WO-US15264.
PR 28-JUL-2000; 2000WO-US20710.
PR 24-AUG-2000; 2000WO-US23328.
PR 08-NOV-2000; 2000WO-US30952.
PR 01-DEC-2000; 2000WO-US32678.

PR 20-DEC-2000; 2000WO-US34956.
PR 28-FEB-2001; 2001WO-US06520.
PR 01-JUN-2001; 2001WO-US17800.
PR 20-JUN-2001; 2001WO-US19692.
PR 29-JUN-2001; 2001WO-US21066.
PR 09-JUL-2001; 2001WO-US21735.
PR 29-AUG-2001; 2001WO-US27099.
PR 18-SEP-1997; 97US-059263P.
PR 18-SEP-1997; 97US-059266P.
PR 17-OCT-1997; 97US-062250P.
PR 21-OCT-1997; 97US-063486P.
PR 24-OCT-1997; 97US-063120P.
PR 24-OCT-1997; 97US-063121P.
PR 28-OCT-1997; 97US-063540P.
PR 28-OCT-1997; 97US-063541P.
PR 28-OCT-1997; 97US-063544P.
PR 28-OCT-1997; 97US-063564P.
PR 29-OCT-1997; 97US-063734P.
PR 31-OCT-1997; 97US-063870P.
PR 31-OCT-1997; 97US-064103P.
PR 13-NOV-1997; 97US-065311P.
PR 21-NOV-1997; 97US-066120P.
PR 24-NOV-1997; 97US-066466P.
PR 24-NOV-1997; 97US-066772P.
PR 11-DEC-1997; 97US-069335P.
PR 12-DEC-1997; 97US-069425P.
PR 17-DEC-1997; 97US-069870P.
PR 18-DEC-1997; 97US-068017P.
PR 10-MAR-1998; 98US-077450P.
PR 11-MAR-1998; 98US-077632P.
PR 11-MAR-1998; 98US-077649P.
PR 20-MAR-1998; 98US-078886P.
PR 20-MAR-1998; 98US-078939P.
PR 27-MAR-1998; 98US-079664P.
PR 27-MAR-1998; 98US-079786P.
PR 31-MAR-1998; 98US-080107P.
PR 31-MAR-1998; 98US-080194P.
PR 01-APR-1998; 98US-080327P.
PR 01-APR-1998; 98US-080333P.
PR 08-APR-1998; 98US-081049P.
PR 08-APR-1998; 98US-081070P.
PR 09-APR-1998; 98US-081195P.
PR 15-APR-1998; 98US-081838P.
PR 21-APR-1998; 98US-082568P.
PR 21-APR-1998; 98US-082569P.
PR 22-APR-1998; 98US-082704P.
PR 22-APR-1998; 98US-082797P.
PR 28-APR-1998; 98US-083322P.
PR 29-APR-1998; 98US-083495P.
PR 29-APR-1998; 98US-083496P.
PR 29-APR-1998; 98US-083499P.
PR 29-APR-1998; 98US-083559P.
PR 05-MAY-1998; 98US-084366P.
PR 06-MAY-1998; 98US-084414P.
PR 07-MAY-1998; 98US-084639P.
PR 07-MAY-1998; 98US-084640P.
PR 07-MAY-1998; 98US-084643P.
PR 15-MAY-1998; 98US-085579P.
PR 15-MAY-1998; 98US-085580P.
PR 15-MAY-1998; 98US-085582P.
PR 15-MAY-1998; 98US-085700P.
PR 18-MAY-1998; 98US-086023P.
PR 22-MAY-1998; 98US-086392P.
PR 22-MAY-1998; 98US-086486P.
PR 28-MAY-1998; 98US-087098P.
PR 28-MAY-1998; 98US-087208P.
PR 02-JUN-1998; 98US-087609P.
PR 02-JUN-1998; 98US-087759P.
PR 03-JUN-1998; 98US-087827P.
PR 04-JUN-1998; 98US-088025P.
PR 04-JUN-1998; 98US-088028P.
PR 04-JUN-1998; 98US-088029P.
PR 04-JUN-1998; 98US-088033P.

PR 04-JUN-1998; 98US-088326P.
PR 05-JUN-1998; 98US-088167P.
PR 05-JUN-1998; 98US-088202P.
PR 05-JUN-1998; 98US-088212P.
PR 05-JUN-1998; 98US-088217P.
PR 09-JUN-1998; 98US-088655P.
PR 10-JUN-1998; 98US-088722P.
PR 10-JUN-1998; 98US-088738P.
PR 10-JUN-1998; 98US-088740P.
PR 10-JUN-1998; 98US-088811P.
PR 10-JUN-1998; 98US-088824P.
PR 10-JUN-1998; 98US-088825P.
PR 10-JUN-1998; 98US-088826P.
PR 11-JUN-1998; 98US-088861P.
PR 11-JUN-1998; 98US-088863P.
PR 11-JUN-1998; 98US-088876P.
PR 12-JUN-1998; 98US-089090P.
PR 12-JUN-1998; 98US-089105P.
PR 16-JUN-1998; 98US-089512P.
PR 16-JUN-1998; 98US-089514P.
PR 17-JUN-1998; 98US-089538P.
PR 17-JUN-1998; 98US-089598P.
PR 17-JUN-1998; 98US-089653P.
PR 18-JUN-1998; 98US-089908P.
PR 19-JUN-1998; 98US-089952P.
PR 22-JUN-1998; 98US-090246P.
PR 22-JUN-1998; 98US-090252P.
PR 22-JUN-1998; 98US-090254P.
PR 24-JUN-1998; 98US-090429P.
PR 24-JUN-1998; 98US-090435P.
PR 24-JUN-1998; 98US-090444P.
PR 24-JUN-1998; 98US-090461P.
PR 24-JUN-1998; 98US-090535P.
PR 24-JUN-1998; 98US-090540P.
PR 25-JUN-1998; 98US-090676P.
PR 25-JUN-1998; 98US-090678P.
PR 25-JUN-1998; 98US-090688P.
PR 25-JUN-1998; 98US-090690P.
PR 25-JUN-1998; 98US-090694P.
PR 25-JUN-1998; 98US-090695P.
PR 25-JUN-1998; 98US-090696P.
PR 26-JUN-1998; 98US-090862P.
PR 26-JUN-1998; 98US-090863P.
PR 26-JUN-1998; 98US-091010P.
PR 01-JUL-1998; 98US-091359P.
PR 01-JUL-1998; 98US-091544P.
PR 02-JUL-1998; 98US-091478P.
PR 02-JUL-1998; 98US-091486P.
PR 02-JUL-1998; 98US-091626P.
PR 02-JUL-1998; 98US-091628P.
PR 02-JUL-1998; 98US-091632P.
PR 24-JUL-1998; 98US-094006P.
PR 04-AUG-1998; 98US-095282P.
PR 10-AUG-1998; 98US-095998P.
PR 10-AUG-1998; 98US-096012P.
PR 17-AUG-1998; 98US-096757P.
PR 17-AUG-1998; 98US-096766P.
PR 17-AUG-1998; 98US-096867P.
PR 17-AUG-1998; 98US-096891P.
PR 17-AUG-1998; 98US-096897P.
PR 18-AUG-1998; 98US-096949P.
PR 18-AUG-1998; 98US-096959P.
PR 18-AUG-1998; 98US-097022P.
PR 26-AUG-1998; 98US-097952P.
PR 26-AUG-1998; 98US-097954P.
PR 26-AUG-1998; 98US-097955P.
PR 26-AUG-1998; 98US-097971P.
PR 26-AUG-1998; 98US-097974P.
PR 26-AUG-1998; 98US-098014P.
PR 01-SEP-1998; 98US-098716P.
PR 01-SEP-1998; 98US-098723P.
PR 02-SEP-1998; 98US-098803P.
PR 02-SEP-1998; 98US-098821P.

PR 02-SEP-1998; 98US-098843P.
PR 09-SEP-1998; 98US-099602P.
PR 10-SEP-1998; 98US-099741P.
Query Match 98.7%; Score 1101; DB 24; Length 382;
Best Local Similarity 96.3%; Pred. No. 9.6e-104;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
QY 1 ATGRLLS-----GQVCRGGTQRPCYKVIYFHDTSRRLNFEFAKEACRRDGGQLVSI 52
Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| 81
QY 53 ESEDEQKLIKFIENLLPSDGDGFWIGLRRRREKQSNSTACQDLYAWTDGSIQFRNWWYVD 112
Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| 141
QY 113 EPSCGSEVCVMYHQPSAPAGIGGPPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGE 172
Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| 201
QY 173 ETELTTPVLPEETQEDAKKTFKESREAAALNLAY 206
Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| 235
RESULT 15
ABU65911
ID ABU65911 standard; Protein; 382 AA.
XX
AC ABU65911;
XX
DT 20-MAY-2003 (first entry)
XX
DE Novel human secreted and transmembrane protein PRO234.
XX
KW Human; secreted protein; transmembrane protein; cytostatic;
KW gene therapy; TNF-Agonist-Alpha; chondrocyte stimulator; tumour;
KW adrenal tumour; lung tumour; colon tumour; breast tumour;
KW prostate tumour; rectal tumour; cervical tumour; liver tumour.
XX
OS Homo sapiens.
XX
PN US2003036157-A1.
XX
PD 20-FEB-2003.
XX
PF 02-JUL-2002; 2002US-0188769.
XX
PR 16-SEP-1998; 98WO-US19330.
PR 07-OCT-1998; 98WO-US21141.
PR 01-DEC-1998; 98WO-US25108.
PR 08-MAR-1999; 99WO-US05028.
PR 14-MAY-1999; 99WO-US10733.
PR 02-JUN-1999; 99WO-US12252.
PR 01-SEP-1999; 99WO-US20111.
PR 15-SEP-1999; 99WO-US21090.
PR 01-DEC-1999; 99WO-US28301.
PR 02-DEC-1999; 99WO-US28551.
PR 30-DEC-1999; 99WO-US31274.
PR 05-JAN-2000; 2000WO-US00219.
PR 18-FEB-2000; 2000WO-US04341.
PR 18-FEB-2000; 2000WO-US04342.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US05004.
PR 01-MAR-2000; 2000WO-US05601.
PR 02-MAR-2000; 2000WO-US05841.
PR 15-MAR-2000; 2000WO-US06884.
PR 30-MAR-2000; 2000WO-US08439.
PR 17-MAY-2000; 2000WO-US13705.
PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.
PR 02-JUN-2000; 2000WO-US15264.
PR 28-JUL-2000; 2000WO-US20710.
PR 24-AUG-2000; 2000WO-US23328.
PR 08-NOV-2000; 2000WO-US30952.
PR 01-DEC-2000; 2000WO-US32678.
PR 20-DEC-2000; 2000WO-US34956.
PR 28-FEB-2001; 2001WO-US06520.
PR 01-JUN-2001; 2001WO-US17800.
PR 20-JUN-2001; 2001WO-US19692.
PR 29-JUN-2001; 2001WO-US21066.
PR 09-JUL-2001; 2001WO-US21735.
PR 29-AUG-2001; 2001WO-US27099.
PR 18-SEP-1997; 97US-059263P.
PR 18-SEP-1997; 97US-059266P.
PR 17-OCT-1997; 97US-062250P.
PR 21-OCT-1997; 97US-063486P.
PR 24-OCT-1997; 97US-063120P.
PR 24-OCT-1997; 97US-063121P.
PR 28-OCT-1997; 97US-063540P.
PR 28-OCT-1997; 97US-063541P.
PR 28-OCT-1997; 97US-063544P.
PR 28-OCT-1997; 97US-063564P.
PR 29-OCT-1997; 97US-063734P.
PR 31-OCT-1997; 97US-063870P.
PR 31-OCT-1997; 97US-064103P.
PR 13-NOV-1997; 97US-065311P.
PR 21-NOV-1997; 97US-066120P.
PR 24-NOV-1997; 97US-066466P.
PR 24-NOV-1997; 97US-066772P.
PR 11-DEC-1997; 97US-069335P.
PR 12-DEC-1997; 97US-069425P.
PR 17-DEC-1997; 97US-069870P.
PR 18-DEC-1997; 97US-068017P.
PR 10-MAR-1998; 98US-077450P.
PR 11-MAR-1998; 98US-077632P.
PR 11-MAR-1998; 98US-077649P.
PR 20-MAR-1998; 98US-078886P.
PR 20-MAR-1998; 98US-078939P.
PR 27-MAR-1998; 98US-079664P.
PR 27-MAR-1998; 98US-079786P.
PR 31-MAR-1998; 98US-080107P.
PR 31-MAR-1998; 98US-080194P.
PR 01-APR-1998; 98US-080327P.
PR 01-APR-1998; 98US-080333P.
PR 08-APR-1998; 98US-081049P.
PR 08-APR-1998; 98US-081070P.
PR 09-APR-1998; 98US-081195P.
PR 15-APR-1998; 98US-081838P.
PR 21-APR-1998; 98US-082568P.
PR 21-APR-1998; 98US-082569P.
PR 22-APR-1998; 98US-082704P.
PR 22-APR-1998; 98US-082797P.
PR 28-APR-1998; 98US-083322P.
PR 29-APR-1998; 98US-083495P.
PR 29-APR-1998; 98US-083496P.
PR 29-APR-1998; 98US-083499P.
PR 29-APR-1998; 98US-083559P.
PR 05-MAY-1998; 98US-084366P.
PR 06-MAY-1998; 98US-084414P.
PR 07-MAY-1998; 98US-084639P.
PR 07-MAY-1998; 98US-084640P.
PR 07-MAY-1998; 98US-084643P.
PR 15-MAY-1998; 98US-085579P.
PR 15-MAY-1998; 98US-085580P.
PR 15-MAY-1998; 98US-085582P.
PR 15-MAY-1998; 98US-085700P.
PR 18-MAY-1998; 98US-086023P.
PR 22-MAY-1998; 98US-086392P.
PR 22-MAY-1998; 98US-086486P.
PR 28-MAY-1998; 98US-087098P.
PR 28-MAY-1998; 98US-087208P.
PR 02-JUN-1998; 98US-087609P.
PR 02-JUN-1998; 98US-087759P.
PR 03-JUN-1998; 98US-087827P.
PR 04-JUN-1998; 98US-088025P.

PR	04-JUN-1998;	98US-088028P;
PR	04-JUN-1998;	98US-088029P;
PR	04-JUN-1998;	98US-088033P;
PR	04-JUN-1998;	98US-088326P;
PR	05-JUN-1998;	98US-088167P;
PR	05-JUN-1998;	98US-088202P;
PR	05-JUN-1998;	98US-088212P;
PR	05-JUN-1998;	98US-088217P;
PR	09-JUN-1998;	98US-088655P;
PR	10-JUN-1998;	98US-088722P;
PR	10-JUN-1998;	98US-088738P;
PR	10-JUN-1998;	98US-088740P;
PR	10-JUN-1998;	98US-088811P;
PR	10-JUN-1998;	98US-088824P;
PR	10-JUN-1998;	98US-088825P;
PR	10-JUN-1998;	98US-088826P;
PR	11-JUN-1998;	98US-088861P;
PR	11-JUN-1998;	98US-088863P;
PR	11-JUN-1998;	98US-088876P;
PR	12-JUN-1998;	98US-089090P;
PR	12-JUN-1998;	98US-089105P;
PR	16-JUN-1998;	98US-089512P;
PR	16-JUN-1998;	98US-089514P;
PR	17-JUN-1998;	98US-089538P;
PR	17-JUN-1998;	98US-089598P;
PR	17-JUN-1998;	98US-089653P;
PR	18-JUN-1998;	98US-089908P;
PR	19-JUN-1998;	98US-089952P;
PR	22-JUN-1998;	98US-090246P;
PR	22-JUN-1998;	98US-090252P;
PR	22-JUN-1998;	98US-090254P;
PR	24-JUN-1998;	98US-090429P;
PR	24-JUN-1998;	98US-090435P;
PR	24-JUN-1998;	98US-090444P;
PR	24-JUN-1998;	98US-090461P;
PR	24-JUN-1998;	98US-090535P;
PR	24-JUN-1998;	98US-090540P;
PR	25-JUN-1998;	98US-090676P;
PR	25-JUN-1998;	98US-090678P;
PR	25-JUN-1998;	98US-090688P;
PR	25-JUN-1998;	98US-090690P;
PR	25-JUN-1998;	98US-090694P;
PR	25-JUN-1998;	98US-090695P;
PR	25-JUN-1998;	98US-090696P;
PR	26-JUN-1998;	98US-090862P;
PR	26-JUN-1998;	98US-090863P;
PR	26-JUN-1998;	98US-091010P;
PR	01-JUL-1998;	98US-091359P;
PR	01-JUL-1998;	98US-091544P;
PR	02-JUL-1998;	98US-091478P;
PR	02-JUL-1998;	98US-091486P;
PR	02-JUL-1998;	98US-091626P;
PR	02-JUL-1998;	98US-091628P;
PR	02-JUL-1998;	98US-091632P;
PR	24-JUL-1998;	98US-094006P;
PR	04-AUG-1998;	98US-095282P;
PR	10-AUG-1998;	98US-095998P;
PR	10-AUG-1998;	98US-096012P;
PR	17-AUG-1998;	98US-096757P;
PR	17-AUG-1998;	98US-096766P;
PR	17-AUG-1998;	98US-096867P;
PR	17-AUG-1998;	98US-096891P;
PR	18-AUG-1998;	98US-096949P;
PR	18-AUG-1998;	98US-096959P;
PR	18-AUG-1998;	98US-097022P;
PR	26-AUG-1998;	98US-097952P;
PR	26-AUG-1998;	98US-097954P;
PR	26-AUG-1998;	98US-097955P;
PR	26-AUG-1998;	98US-097971P;
PR	26-AUG-1998;	98US-097974P;
PR	26-AUG-1998;	98US-098014P;
PR	01-SEP-1998;	98US-098716P;

PR	01-SEP-1998;	98US-098723P.
PR	02-SEP-1998;	98US-098803P.
PR	02-SEP-1998;	98US-098821P.
PR	02-SEP-1998;	98US-098843P.
PR	09-SEP-1998;	98US-099602P.
PR	10-SEP-1998;	98US-099741P.
PR	10-SEP-1998;	98US-099754P.
PR	10-SEP-1998;	98US-099763P.

Query Match	98.7%	Score 1101;	DB 24;	Length 382;
Best Local Similarity	96.3%	Pred. No. 9.6e-104;		
Matches 206; Conservative	0;	Mismatches 0;	Indels 8;	Gaps 1;

QY	1	ATGRLLS-----GQVPCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI	52
Db	22	ATGRLLSASDLDLRGGQPVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI	81
QY	53	ESEDEQKLIEKFIENLLPSDGDGFWIGLRRREEKQSNSTACQDLVATWDGSI SQFRNWWYVD	112
Db	82	ESEDEQKLIEKFIENLLPSDGDGFWIGLRRREEKQSNSTACQDLVATWDGSI SQFRNWWYVD	141
QY	113	EPSCGSEVCVVMYHQPSPAGIGGYPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGE	172
Db	142	EPSCGSEVCVVMYHQPSPAGIGGYPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGE	201
QY	173	ETELTTPVLPEETQEEADAKKTFKESREAAALNLAY	206
Db	202	ETELTTPVLPEETQEEADAKKTFKESREAAALNLAY	235

Search completed: December 22, 2003, 16:10:07
Job time : 30.1819 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 22, 2003, 16:05:39 ; Search time 43.5759 Seconds
(without alignments)
2214.797 Million cell updates/sec

Title: US-09-887-855-2
Perfect score: 2000
Sequence: 1 MRPGTALQAVLLAVLLVGLR.....PDQMGRSKESGWENEIYGY 374

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : SPTREMBL 23:*
- 1: sp_archaea:*
 - 2: sp_bacteria:*
 - 3: sp_fungi:*
 - 4: sp_human:*
 - 5: sp_invertebrate:*
 - 6: sp_mammal:*
 - 7: sp_mhc:*
 - 8: sp_organelle:*
 - 9: sp_phage:*
 - 10: sp_plant:*
 - 11: sp_rodent:*
 - 12: sp_virus:*
 - 13: sp_vertebrate:*
 - 14: sp_unclassified:*
 - 15: sp_rvirus:*
 - 16: sp_bacteriap:*
 - 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2000	100.0	374	4	Q96NF3
2	1997	99.9	374	4	Q8TAY8
3	1996	99.8	374	4	Q96NC5
4	1697.5	84.9	374	11	Q9Z209
5	862.5	43.1	211	11	Q8C351
6	652	32.6	292	11	Q8BVU2
7	646	32.3	246	11	Q8BMT7
8	185	9.2	1290	13	Q9W6E1
9	178.5	8.9	1456	11	Q61830
10	177.5	8.9	1348	5	Q25199
11	176.5	8.8	1479	4	Q9Y5P9
12	176.5	8.8	1479	4	Q9UBG0
13	170.5	8.5	1479	11	Q64449
14	169	8.5	217	11	Q8C4F8
15	165.5	8.3	742	11	Q8K4Q8
16	165.5	8.3	742	11	Q8C979

17	161.5	8.1	134	5	Q9XYX3	Q9xyx3 hydra magni
18	159.5	8.0	742	11	Q8VIF6	Q8vif6 mus musculu
19	159	8.0	1031	5	Q8WSX2	Q8wsx2 dugesia tig
20	158.5	7.9	719	6	O62623	O62623 bos taurus
21	158	7.9	142	11	Q8CJ86	Q8cj86 mus musculu
22	158	7.9	142	11	Q8BHK7	Q8bhk7 mus musculu
23	158	7.9	295	11	Q91ZM4	Q91zm4 mus musculu
24	158	7.9	311	11	Q9D8V4	Q9d8v4 mus musculu
25	158	7.9	325	11	Q91ZX0	Q91zx0 mus musculu
26	157.5	7.9	358	6	Q8HY04	Q8hy04 papio hamad
27	157.5	7.9	381	6	Q8SQB2	Q8sqb2 macaca mula
28	157.5	7.9	652	4	Q8IXK1	Q8ixk1 homo sapien
29	157	7.8	158	13	Q90WI7	Q90wi7 bungarus fa
30	156	7.8	162	5	Q25459	Q25459 megabalanus
31	156	7.8	323	11	Q8CJ91	Q8cj91 mus musculu
32	156	7.8	339	6	Q95244	Q95244 sus scrofa
33	155	7.8	293	11	Q8BGZ0	Q8bgz0 mus musculu
34	155	7.8	323	11	Q8CJ94	Q8cj94 mus musculu
35	155	7.8	323	11	Q8CJ93	Q8cj93 mus musculu
36	155	7.8	323	11	Q8CJ88	Q8cj88 mus musculu
37	155	7.8	381	6	Q95LA8	Q95la8 macaca mula
38	155	7.8	1152	13	Q90WM2	Q90wm2 xenopus lae
39	154	7.7	312	6	Q8HXL6	Q8hxl6 macaca mula
40	154	7.7	322	11	Q8CJ89	Q8cj89 mus musculu
41	154	7.7	323	11	Q8CJ92	Q8cj92 mus musculu
42	154	7.7	381	6	Q95LC6	Q95lc6 macaca neme
43	154	7.7	381	6	Q8HXL7	Q8hxl7 macaca mula
44	154	7.7	404	6	Q95J96	Q95j96 macaca mula
45	152.5	7.6	427	6	Q8HYB9	Q8hyb9 pan troglod

ALIGNMENTS

RESULT 1

Q96NF3	PRELIMINARY;	PRT;	374 AA.
AC	Q96NF3;		
DT	01-DEC-2001 (Tremblrel. 19, Created)		
DT	01-DEC-2001 (Tremblrel. 19, Last sequence update)		
DT	01-MAR-2003 (Tremblrel. 23, Last annotation update)		
DE	Hypothetical protein FLJ30977.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Ishibashi T., Kanehori K., Yosida M., Watanabe S., Ishida S., Ono Y.,		
RA	Hotuta T., Hiraoka S., Murakawa K., Takiguchi S., Kusano J.,		
RA	Watanabe M., Fujimori K., Tanai H., Ishida M., Yamashita H., Chiba Y.,		
RA	Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S.,		
RA	Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,		
RA	Kimura K., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,		
RA	Wagatsuma M., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,		
RA	Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;		
RT	"NEDO human cDNA sequencing project."		
RL	Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AK055539; BAB70946.1; ..		
DR	InterPro; IPR001304; Lectin_C.		
DR	Pfam; PF00059; lectin c; 1.		
DR	SMART; SM00034; CLECT-1.....		
DR	PROSITE; PS50041; C TYPE LECTIN_2; 1.		
KW	Hypothetical protein.		
SQ	SEQUENCE 374 AA; 42280 MW; 8AE64E5BC9E56DCD CRC64;		

Query Match 100.0%; Score 2000; DB 4; Length 374;
Best Local Similarity 100.0%; Pred. No. 1.5e-180;
Matches 374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRCPCYKVIYFHTSRRLNFEAK	60
Db	1	MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRCPCYKVIYFHTSRRLNFEAK	60

QY 61 EACRRDGGQLVSI ESEDEQKLI EKFIENLLPSDGDGFWIGLRRRREKQSNSTACQDLYAWT 120
Db 61 EACRRDGGQLVSI ESEDEQKLI EKFIENLLPSDGDGFWIGLRRRREKQSNSTACQDLYAWT 120
QY 121 DGSISQFRNWYVDEPSCGSEVCVVMYHQP SAPAGIGGYPYMFQWDDRCNMKNFICKYSD 180
Db 121 DGSISQFRNWYVDEPSCGSEVCVVMYHQP SAPAGIGGYPYMFQWDDRCNMKNFICKYSD 180
QY 181 EKPAVPSREAEGEETELTPVLPEETQEEADAKKTFKESREAAALNLAYILIPSIPLLLLLV 240
Db 181 EKPAVPSREAEGEETELTPVLPEETQEEADAKKTFKESREAAALNLAYILIPSIPLLLLLV 240
QY 241 VTTVVCWWVICRKRKRQPDPSSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300
Db 241 VTTVVCWWVICRKRKRQPDPSSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300
QY 301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360
Db 301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360
QY 361 SKESGWVENEIYGY 374
Db 361 SKESGWVENEIYGY 374

RESULT 2

Q8TAY8 PRELIMINARY; PRT; 374 AA.
AC Q8TAY8;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Similar to unnamed protein product.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC025407; AAH25407.1; -.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00059; lectin_c; 1.
DR SMART; SM00034; CLECT_1.
DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
SQ SEQUENCE 374 AA; 42312 MW; FC214E6BC9E578D9 CRC64;

Query Match 99.9%; Score 1997; DB 4; Length 374;
Best Local Similarity 99.7%; Pred. No. 2.9e-180;
Matches 373; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEAK 60
Db 1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEAK 60
QY 61 EACRRDGGQLVSI ESEDEQKLI EKFIENLLPSDGDGFWIGLRRRREKQSNSTACQDLYAWT 120
Db 61 EACRRDGGQLVSI ESEDEQKLI EKFIENLLPSDGDGFWIGLRRRREKQSNSTACQDLYAWT 120
QY 121 DGSISQFRNWYVDEPSCGSEVCVVMYHQP SAPAGIGGYPYMFQWDDRCNMKNFICKYSD 180
Db 121 DGSISQFRNWYVDEPSCGSEVCVVMYHQP SAPAGIGGYPYMFQWDDRCNMKNFICKYSD 180
QY 181 EKPAVPSREAEGEETELTPVLPEETQEEADAKKTFKESREAAALNLAYILIPSIPLLLLLV 240
Db 181 EKPAVPSREAEGEETELTPVLPEETQEEADAKKTFKESREAAALNLAYILIPSIPLLLLLV 240
QY 241 VTTVVCWWVICRKRKRQPDPSSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300
Db 241 VTTVVCWWVICRKRKRQPDPSSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300

QY 301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360
Db 301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360
QY 361 SKESGWVENEIYGY 374
Db 361 SKESGWVENEIYGY 374

RESULT 3

Q96NC5 PRELIMINARY; PRT; 374 AA.
AC Q96NC5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein FLJ31092.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,
RA Fukuzumi Y., Fujimori Y., Komiyama M., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,
RA Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A.,
RA Kawakami B., Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K.,
RA Isogai T.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK055654; BAB70978.1; -.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00059; lectin_c; 1.
DR SMART; SM00034; CLECT_1.
DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW Hypothetical protein.
SQ SEQUENCE 374 AA; 42310 MW; CBF74E676E23BA16 CRC64;

Query Match 99.8%; Score 1996; DB 4; Length 374;
Best Local Similarity 99.7%; Pred. No. 3.7e-180;
Matches 373; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEAK 60
Db 1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEAK 60
QY 61 EACRRDGGQLVSI ESEDEQKLI EKFIENLLPSDGDGFWIGLRRRREKQSNSTACQDLYAWT 120
Db 61 EACRRDGGQLVSI ESEDEQKLI EKFIENLLPSDGDGFWIGLRRRREKQSNSTACQDLYAWT 120
QY 121 DGSISQFRNWYVDEPSCGSEVCVVMYHQP SAPAGIGGYPYMFQWDDRCNMKNFICKYSD 180
Db 121 DGSISQFRNWYVDEPSCGSEVCVVMYHQP SAPAGIGGYPYMFQWDDRCNMKNFICKYSD 180
QY 181 EKPAVPSREAEGEETELTPVLPEETQEEADAKKTFKESREAAALNLAYILIPSIPLLLLLV 240
Db 181 EKPAVPSREAEGEETELTPVLPEETQEEADAKKTFKESREAAALNLAYILIPSIPLLLLLV 240
QY 241 VTTVVCWWVICRKRKRQPDPSSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300
Db 241 VTTVVCWWVICRKRKRQPDPSSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300
QY 301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360
Db 301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGR 360
QY 361 SKESGWVENEIYGY 374
Db 361 SKESGWVENEIYGY 374

```
RESULT 4
Q9Z209 PRELIMINARY; PRT; 374 AA.
AC Q9Z209;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Layilin.
OS Cricetulus griseus (Chinese hamster).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Cricetulus.
OX NCBI_TaxID=10029;
RN [1]
RP SEQUENCE FROM N.A.
RA Borowsky M.L., Hynes R.O.;
RT "Layilin, a novel talin-binding transmembrane protein homologous with
RT C-type lectins, is localized in membrane ruffles.";
RL J. Cell Biol. 143:0-0(1998)
DR EMBL; AF093673; AAC68695.1;
DR HSSP; P22897; 1EGG.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00059; lectin_c; 1.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS0041; C_TYPE_LLECTIN 2; 1.
SQ SEQUENCE 374 AA; 42435 MW; 298A8BA24FB04E1C CRC64;

Query Match 84.9%; Score 1697.5; DB 11; Length 374;
Best Local Similarity 84.7%; Pred. No. 6.2e-152;
Matches 316; Conservative 23; Mismatches 33; Indels 1; Gaps 1;

QY 1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQPCYKVIYFHDTSRRLNFEAK 60
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1 MQPGPALQAVLLAVLLSEPRSSKGRLLSGQLVCRGGTTRPCYKVIYFHDAFQRLNFEAK 60

QY 61 EACRRDGGQLVLSIESEDEQKLIKFIENLLPSDGFWDGFWIGLRRRREKQSNSTACQDLYAWT 120
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
61 EACRRDGGQLVLSIETEDQRLIEKFIENLLASDGDGFWIGLRRLEVKQVNNACQDLYAWT 120

QY 121 DGSISQFRNWWYVDEPSCGSEVVMYHQPAPAGIGGPMFQWDDRCNMKNPFICKYSYD 180
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
121 DGSTSQFRNWWYVDEPSCGSEVVMYHQPAPPGIGGPMFQWDDRCNMKNPFICKYAD 180

QY 181 EKPA-VPSREAEGETELTPVPEETQEDAKTKPKESREAAALNLAYILIPSLILL 239
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
181 EKPSTPSIRPGGEATEPPTVPEETQKEDTKETPKESREAAALNLAYILIPSLILL 240

QY 240 VVTVCVWVICRKRKQDPDPSTKKOHTIWPSPHQNSPDLEVNVIRKQSEADLAETR 299
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
241 VVTSAAACVWICRRRKQEQPDPTTKEHTIWPTHQNSPNLDVNVIRKQSEADLTEPR 300

QY 300 PDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVEGFTNDIYFSPDMG 359
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
301 PDLKNISFRVCSSEAPPDDISCDYDNMAVNPSESGFVTLASMESGFVTDIYFSPDRMG 360

QY 360 RSKESGWENEIY 372
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
361 RSKESGWENEIY 373

RESULT 5
Q8C351 PRELIMINARY; PRT; 211 AA.
AC Q8C351;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE LAYILIN homolog (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

QY 10 VLLAVLLVGLRAATGRLLSGQPVCRGGTQPCYKVIYFHDTSRRLNFEAEACRRDGGQ 69
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
8 LLGAALLCAQGAFAFARRVWSGQKVCFAFADVHKPCYKMAFHELSRVSFQEARLACESEGGV 67

QY 70 LVSIESEDEQKLIKFIENLLP-----SDGDFWIGLRRRREKQSNSTACQDLYAWTDGSI 124
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
68 LLSLENEAEQKLIESMLQNLTKPGTGISDGDGFWIGLRLSGDGT-SGACPDLYQWSDGSS 126

QY 125 SQFRNWWYVDEPSCGSEVVMYHQPAPAGIGGPMFQWDDRCNMKNPFICKYSDE-KP 183
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
|
```

```
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Lung;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK086930; BAC39765.1;
FT NON TER 211
SQ SEQUENCE 211 AA; 23697 MW; AD9870B5957DD5AE CRC64;

Query Match 43.1%; Score 862.5; DB 11; Length 211;
Best Local Similarity 77.7%; Pred. No. 2.6e-73;
Matches 164; Conservative 12; Mismatches 26; Indels 9; Gaps 2;

QY 1 MRPGTALQAVLLAVLLVGLRAATGRLLS-----GQPVCRGGTQPCYKVIYFHDTSR 52
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1 MQPGAALQAVLLAVLLAKPRDSKGRLLSASDLDPGGQLVCRGGTTRPCYKVIYFHDAFQ 60

QY 53 RLNFEEAKEACRRDGGQLVLSIESEDEQKLIKFIENLLPSDGFWDGFWIGLRRRREKQSNSTA 112
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
61 RLNFEEAKETCMEDGGQLVSIETEDQRLIEKFIENLLASDGDGFWIGLKRLEEKQSNNTA 120

QY 113 CQDLYAWTDGSIQFRNWWYVDEPSCGSEVVMYHQPAPAGIGGPMFQWDDRCNMKN 172
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
121 CQDLYAWTDGSTSQFRNWWYVDEPSCGSEVVMYHQPAPPGIGGPMFQWDDRCNMKN 180

QY 173 NFICKYSDEKPA-VPSREAEGETELTPVL 202
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
181 NFICKYHDDKPSITTPSINPGGEATEPATPLL 211

RESULT 6
Q8BVU2 PRELIMINARY; PRT; 292 AA.
ID Q8BVU2;
AC Q8BVU2;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE C-type lectin protein MT75 homolog.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK076523; BAC36378.1;
SQ SEQUENCE 292 AA; 32502 MW; 73E631C0714D54E2 CRC64;

Query Match 32.6%; Score 652; DB 11; Length 292;
Best Local Similarity 47.0%; Pred. No. 3.2e-53;
Matches 140; Conservative 39; Mismatches 79; Indels 40; Gaps 9;

QY 10 VLLAVLLVGLRAATGRLLSGQPVCRGGTQPCYKVIYFHDTSRRLNFEAEACRRDGGQ 69
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
8 LLGAALLCAQGAFAFARRVWSGQKVCFAFADVHKPCYKMAFHELSRVSFQEARLACESEGGV 67

QY 70 LVSIESEDEQKLIKFIENLLP-----SDGDFWIGLRRRREKQSNSTACQDLYAWTDGSI 124
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
68 LLSLENEAEQKLIESMLQNLTKPGTGISDGDGFWIGLRLSGDGT-SGACPDLYQWSDGSS 126

QY 125 SQFRNWWYVDEPSCGSEVVMYHQPAPAGIGGPMFQWDDRCNMKNPFICKYSDE-KP 183
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
|
```


Db 127 SQFRNWTDEPSCGSEKCVVMYHQPTANPGLGGPYLYQWNDDRCNMKNHNYICKYEPIHP 186

Qy 184 AVPSRAEAGEETELTPVLPEETOEDAKTKFKESREAL--NLAVILIPSIPLLLLLVV 241

Db 187 TEPA-----EKPYLTNQ--PEETHENVV-----VTEAGIIPNLIVYIPTIPLLLILV 233

Qy 242 TTVCVVMWICRKRK-----EQDPSTKK-----QHTIWPSPHQGNSPDL 281

Db 234 ALGTCCFQMLHKKRARRHFIDKSTPLSSECLAEKSLNSLVHMAGSLIPYHFQNNSPSL 291

RESULT 7

Q8BMI7 PRELIMINARY; PRT; 246 AA.

AC Q8BMI7;

DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE C-type lectin protein MT75 homolog.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Forelimb;

RX MEDLINE=22354683; PubMed=12466851;

RA The FANTOM Consortium,

RA the RIKEN Genome Exploration Research Group Phase I & II Team;

RT "Analysis of the mouse transcriptome based on functional annotation of

RT 60,770 full-length cDNAs";

RL Nature 420:563-573(2002).

DR EMBL; AK031063; BAC27234.1; ..

SQ SEQUENCE 246 AA; 27394 MW; E872660C58267752 CRC64;

Query Match 32.3%; Score 646; DB 11; Length 246;

Best Local Similarity 52.2%; Pred. No. 9.5e-53;

Matches 132; Conservative 34; Mismatches 65; Indels 22; Gaps 7;

Qy 10 VLLAVLLVGLRAATGRLLSGQPVCRCGGTQPCYKVIYFHDTSRRLNFEAKEACRRDGGQ 69

Db 8 LLGAALLCAQGAPARRVWSGQKVCFADVHKPCYKWAYFHELSSRVSFQEARLACESEGGV 67

Qy 70 LVSISEDEQKLIKFIENLLP-----SDGDFWIGLRRREEKQSNSTACQDLYAWTDGSI 124

Db 68 LLSLENEAEQKLIESMLQNLTKPGTIGSDGFWIGLRSRGDQT-SGACPDLYQWSDGSS 126

Qy 125 SQFRNWTDEPSCGSEKCVVMYHQPSPAGIGCPYMFQWNDDRCNMKNHNYICKYSE-KP 183

Db 127 SQFRNWTDEPSCGSEKCVVMYHQPTANPGLGGPYLYQWNDDRCNMKNHNYICKYEPIHP 186

Qy 184 AVPSRAEAGEETELTPVLPEETOEDAKTKFKESREAL--NLAVILIPSIPLLLLLVV 241

Db 187 TEPA-----EKPYLTNQ--PEETHENVV-----VTEAGIIPNLIVYIPTIPLLLILV 233

Qy 242 TTVCVVMWICRKR 254

Db 234 ALGTCCFQMLHKKR 246

RESULT 8

Q9W6E1 PRELIMINARY; PRT; 1290 AA.

AC Q9W6E1;

DT 01-NOV-1999 (TrEMBLrel. 12, Created)

DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Neurocan core protein.

OS Gallus gallus (Chicken).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

OC Gallus.

OX NCBI_TaxID=9031;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=20309833; PubMed=10851024;

RA Li H., Leung T.C., Hoffman S., Balsamo J., Lilien J.;

RT "Coordinate Regulation of Cadherin and Integrin Function by the

RT Chondroitin Sulfate Proteoglycan Neurocan.";

RL J. Cell Biol. 149:1275-1288(2000).

DR EMBL; AF116856; AAD24546.2; ..

DR HSSP; P08709; 1BF9.

DR InterPro; IPR000152; Asx_hydroxyl.

DR InterPro; IPR000742; EGF_2.

DR InterPro; IPR001881; EGF_Ca.

DR InterPro; IPR001438; EGF_II.

DR InterPro; IPR006209; EGF_like.

DR InterPro; IPR003599; Ig.

DR InterPro; IPR007110; Ig-like.

DR InterPro; IPR003006; Ig_MHC.

DR InterPro; IPR001304; Lectin_C.

DR InterPro; IPR000538; Link.

DR InterPro; IPR000436; Sushi_SCR_CCP.

DR Pfam; PF00008; EGF; 2.

DR Pfam; PF00047; ig; 1.

DR Pfam; PF00059; lectin_c; 1.

DR Pfam; PF00084; sushi; 1.

DR Pfam; PF00193; Xlink; 2.

DR PRINTS; PR00010; EGFBL00D.

DR PRINTS; PR01265; LINKMODULE.

DR ProDom; PD000918; Link; 2.

DR SMART; SM00032; CCP; 1.

DR SMART; SM00034; CLECT; 1.

DR SMART; SM00179; EGF_CA; 1.

DR SMART; SM00409; IG; 1.

DR SMART; SM00445; LINK; 2.

DR PROSITE; PS00010; ASX_HYDROXYL; 1.

DR PROSITE; PS00615; C_TYPE_LECTIN_1; 1.

DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.

DR PROSITE; PS00022; EGF_1; 2.

DR PROSITE; PS01186; EGF_2; 1.

DR PROSITE; PS01187; EGF_CA; 1.

DR PROSITE; PS50835; IG_LIKE; 1.

DR PROSITE; PS01241; LINK; 2.

KW EGF-like domain.

SQ SEQUENCE 1290 AA; 138877 MW; 182BD86D0E40BE78 CRC64;

Query Match 9.2%; Score 185; DB 13; Length 1290;

Best Local Similarity 34.7%; Pred. No. 3.3e-08;

Matches 50; Conservative 18; Mismatches 42; Indels 34; Gaps 8;

Qy 38 QRPCYKVIYFHDTSRRLNFEAKEACRRDGGQLVSISEDEQKLIKFIENLLPSDGF 97

Db 1064 QGHCYR--YF---SRRRSWEDAERDCRRRAGHLTSHSQBEHGFINSF-----GHENTW 1112

Qy 98 IGLRRREEKQSNSTACQDLYAWTDGSIQFRNWTDEPS---CGSEVGVVY-HQPSAPA 153

Db 1113 IGLNDRIVEQD-----FQWTDNTGLQYENWRENQDPNPFAGGDCVVLVSHE----- 1159

Qy 154 GIGGYPYMFQWNDDRCNMKNHNYICK 177

Db 1160 -IG-----KWNDVPCYNLPYICK 1177

RESULT 9

Q61830 PRELIMINARY; PRT; 1456 AA.

ID Q61830

AC Q61830;

DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Macrophage mannose receptor precursor.

GN MRC1.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```

OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RX MEDLINE=93043353; PubMed=1421407;
RA Harris N., Rits M., Chang G., Ezekowitz R.B.;
RT "Characterization of the murine macrophage mannose receptor.";
RL Blood 80:2363-2373(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RA Super M.;
RL Submitted (APR-1992) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z11974; CAA78028.1; -.
DR HSSP; P22897; 1EGG.
DR MGD; MGI:97142; Mrc1.
DR InterPro; IPR002353; AntifreezeII.
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR000772; Ricin_B_lectin.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00059; lectin_c; 8.
DR Pfam; PF00652; Ricin_B_lectin; 2.
DR PRINTS; PR00356; ANTIFREEZEII.
DR PRINTS; PR00013; FNTYPEII.
DR ProDom; PD000995; FN_Type_II; 1.
DR SMART; SM00034; CLECT; 8.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00458; RICIN; 1.
DR PROSITE; PS00615; C_TYPE_LECTIN_1; 6.
DR PROSITE; PS50041; C_TYPE_LECTIN_2; 8.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS50231; RICIN_B_LECTIN; 1.
KW Receptor; Signal.
FT SIGNAL 1 18 POTENTIAL.
FT CHAIN 19 1456 MACROPHAGE MANNOSE RECEPTOR.
SQ SEQUENCE 1456 AA; 165065 MW; 4EBD3F1B8619A594 CRC64;

Query Match 8.9%; Score 178.5; DB 11; Length 1456;
Best Local Similarity 25.9%; Pred. No. 1.6e-07;
Matches 53; Conservative 33; Mismatches 66; Indels 53; Gaps 8;

QY 42 YKVIYFHDTSRLNFEEAKEACRRDGGQLVSIIESEDEQKLEKFIENLLPSDGDFTWGLR 101
Db 807 YKDYQYFYSKEKETMDNARRFCCKNFGDLATIKSESEKFLWKYI-NKNGGQSPYFIGML 865

QY 102 RREEKQSNSTACQDLVATDGSISQFRNWWYDEPSCGS--EVCVMYHQPSAPAGIGGPY 159
Db 866 ISMDKK-----FIWMDGSKYDFVAVATGEPNFANDDENCVTMY---TNSGF---- 908

QY 160 MFQWDDRCNMKNFICK---YSDEKPAVPSREAEGETELTTPVLPEETQE----- 208
Db 909 ---WNDINGYPNNFICQHNSSINATAMP-----TTPTPGGCKEGWHLYKNK 954

QY 209 -----EDAKTFKESREAAALNL 225
Db 955 CFKIFGFANEKKSWQDARQACKGL 979

RESULT 10
Q25199 PRELIMINARY; PRT; 1348 AA.
AC Q25199;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Tyrosine kinase receptor.
OS Hydra attenuata (Hydra) (Hydra vulgaris).
OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Anthomedusae;
OC Hydridae; Hydra.
OX NCBI_TaxID=6087;
RN [1]
RP SEQUENCE FROM N.A.
```

```

RC STRAIN=Irvine;
RX MEDLINE=20209407; PubMed=10744720;
RA Reidling J.C., Miller M.A., Steele R.E.;
RT "Sweet Tooth, a Novel Receptor Protein-tyrosine Kinase with C-type
RT Lectin-like Extracellular Domains.";
RL J. Biol. Chem. 275:10323-10330(2000).
DR EMBL; L22612; AAA29218.2; -.
DR HSSP; P11362; 1FGK.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR003990; Pancreatib_ac.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00059; lectin_c; 4.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR01504; PNCREATITSAP.
DR ProDom; PD000001; Prot_kinase; 2.
DR SMART; SM00034; CLECT; 4.
DR SMART; SM00219; TyrKC; 1.
DR PROSITE; PS00615; C_TYPE_LECTIN_1; 2.
DR PROSITE; PS50041; C_TYPE_LECTIN_2; 4.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
KW ATP-binding; Kinase; Receptor; Transferase; Tyrosine-protein kinase.
SQ SEQUENCE 1348 AA; 156916 MW; 77D2122093227FEF CRC64;

Query Match 8.9%; Score 177.5; DB 5; Length 1348;
Best Local Similarity 28.0%; Pred. No. 1.8e-07;
Matches 47; Conservative 32; Mismatches 46; Indels 43; Gaps 8;

QY 27 LSGQPVC--RGGTQRPC-----YKVIYFHDTSRLNFEEAKEACRRDGGQLVSIIESE 78
Db 412 LSHRFICKVKRATNEYCAEGWTSYRIYCYFIYSIEFDWPKSFSSCQINGNLLSIENQEE 471

QY 79 QKLEKFIENLLPSDGD-FWIGLRR-----RREEKQSNSTACQDLVATDGSISQFRNWWY 131
Db 472 ----NRFIENDLIKNDKYWIGLKNKIWDYLNKKNR-----FEWSDNTYTQFFNWI 518

QY 132 VDEP--SCGSEVCVMYHQPSAPAGIGGPYMFQWDDRCNMKNFICK 177
Db 519 TNQPDNNNGIESCVEMYN-----GWSDKCKVLNGFICK 553

RESULT 11
Q9Y5P9 PRELIMINARY; PRT; 1479 AA.
AC Q9Y5P9;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Endocytic receptor Endo180.
GN ENDO180.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20148849; PubMed=10683150;
RA Sheikh H., Yarwood H., Ashworth A., Isacke C.;
RT "Endo180, an endocytic recycling glycoprotein related to the
RT macrophage mannose receptor is expressed on fibroblasts, endothelial
RT cells and macrophages and functions as a lectin receptor.";
RL J. Cell Sci. 113:1021-1032(2000).
DR EMBL; AF134838; AAD30280.1; -.
DR HSSP; P02751; 2FN2.
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR000566; Lipocln_cytFABP.
DR InterPro; IPR000772; Ricin_B_lectin.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00059; lectin_c; 8.
DR PRINTS; PR00013; FNTYPEII.
```

DR ProDom; PD000995; FN Type_II; 1.
DR SMART; SM00034; CLECT; 8.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00458; RICIN; 1.
DR PROSITE; PS00615; C_TYPE LECTIN_1; 3.
DR PROSITE; PS50041; C_TYPE LECTIN_2; 8.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00213; LIPOCALIN; 1.
DR PROSITE; PS50231; RICIN_B LECTIN; 1.
KW Receptor.
SQ SEQUENCE 1479 AA; 166669 MW; 9F4BAF355F036FCE CRC64;

Query Match 8.8%; Score 176.5; DB 4; Length 1479;
Best Local Similarity 30.4%; Pred. No. 2.6e-07;
Matches 51; Conservative 25; Mismatches 57; Indels 35; Gaps 7;

QY 28 SQPVCRCGGTQPCYKVIYFHDTSRRLNFEAEKACRRDGGQLVSI ESEDEQKLI EKPIE 87
Db 385 SWQPF-----QGHCYRL-----QAEKRSWQESKKACLRGGDLVSIHSMAELEFITKQIK 434

QY 88 NLLPSDGDGFWIGLRRRREKQSNSTACQDLYAWTDGSI SQFRNWWYVDEPS---CGSEVVCVW 144
Db 435 QEVE---ELWIGL-----NDLKLQMFWSGSLVSFTHWHPFEPNFRDSDL EDCVT 483

QY 145 MYHQPSAPAGIGGPGYMFQWNDRCNMKNFNICKYSDEKPAVPSREAE G 192
Db 484 IW----GPEG-----RWNDSPCNQSLPSICKKAGQLSQGAAEEDHG 520

RESULT 12
Q9UBGO
ID Q9UBGO PRELIMINARY; PRT; 1479 AA.
AC Q9UBGO;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Urokinase receptor-associated protein UPARAP.
GN KIAA0709.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Behrendt N., Jensen O.N., Engelholm L.H., Mortz E., Mann M., Dano K.;
RT "A urokinase receptor-associated protein with specific collagen-binding properties."
RT binding properties."
RL Submitted (NOV-1998) to the EMBL/GenBank/DBSJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=98403880; PubMed=9734811;
RA Ishikawa K., Nagase T., Suyama M., Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RA "Prediction of the coding sequences of unidentified human genes. X. The complete sequences of 100 new cDNA clones from brain which can code for large proteins in vitro."
RL DNA Res. 5:169-176(1998).
DR EMBL; AF107292; AAF14192.1; -.
DR EMBL; AB014609; BAA31684.1; -.
DR HSSP; P02751; 2FN2.
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR000566; Lipocln_cytFABP.
DR InterPro; IPR000772; Ricin_B_lectin.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00059; lectin_c; 8.
DR PRINTS; PR00013; FNTYPEII.
DR ProDom; PD000995; FN_Type_II; 1.
DR SMART; SM00034; CLECT; 8.
DR SMART; SM00059; FN2; 1.
DR PROSITE; PS00615; C_TYPE LECTIN_1; 3.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00213; LIPOCALIN; 1.
DR PROSITE; PS50231; RICIN_B LECTIN; 1.
KW Receptor.
SQ SEQUENCE 1479 AA; 166669 MW; 9F4BAF355F036FCE CRC64;

Query Match 8.8%; Score 176.5; DB 4; Length 1479;
Best Local Similarity 30.4%; Pred. No. 2.6e-07;
Matches 51; Conservative 25; Mismatches 57; Indels 35; Gaps 7;

QY 28 SQPVCRCGGTQPCYKVIYFHDTSRRLNFEAEKACRRDGGQLVSI ESEDEQKLI EKPIE 87
Db 385 SWQPF-----QGHCYRL-----QAEKRSWQESKKACLRGGDLVSIHSMAELEFITKQIK 434

QY 88 NLLPSDGDGFWIGLRRRREKQSNSTACQDLYAWTDGSI SQFRNWWYVDEPS---CGSEVVCVW 144
Db 435 QEVE---ELWIGL-----NDLKLQMFWSGSLVSFTHWHPFEPNFRDSDL EDCVT 483

QY 145 MYHQPSAPAGIGGPGYMFQWNDRCNMKNFNICKYSDEKPAVPSREAE G 192
Db 484 IW----GPEG-----RWNDSPCNQSLPSICKKAGQLSQGAAEEDHG 520

RESULT 13
Q64449
ID Q64449 PRELIMINARY; PRT; 1479 AA.
AC Q64449;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Lectin lambda.
GN MRC2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96355501; PubMed=8702911;
RA Wu K., Yuan J., Lasky L.A.;
RT "Characterization of a novel member of the macrophage mannose receptor type C lectin family."
RL J. Biol. Chem. 271:21323-21330(1996).
DR EMBL; U56734; AAC52729.1; -.
DR HSSP; P02751; 2FN2.
DR MGD; MGI:107818; Mrc2.
DR InterPro; IPR002353; AntifreezeII.
DR InterPro; IPR001128; Cytochrome P450.
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR000566; Lipocln_cytFABP.
DR InterPro; IPR000772; Ricin_B_lectin.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00059; lectin_c; 8.
DR PRINTS; PR00356; ANTIFREEZEII.
DR PRINTS; PR00013; FNTYPEII.
DR ProDom; PD000995; FN_Type_II; 1.
DR SMART; SM00034; CLECT; 8.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00458; RICIN; 1.
DR PROSITE; PS00086; CYTOCHROME P450; 1.
DR PROSITE; PS00615; C_TYPE LECTIN_1; 3.
DR PROSITE; PS50041; C_TYPE LECTIN_2; 8.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00213; LIPOCALIN; 1.
DR PROSITE; PS50231; RICIN_B LECTIN; 1.
SQ SEQUENCE 1479 AA; 167112 MW; 62D456E10B9B48C1 CRC64;

Query Match 8.5%; Score 170.5; DB 11; Length 1479;
Best Local Similarity 31.4%; Pred. No. 9.4e-07;
Matches 48; Conservative 22; Mismatches 48; Indels 35; Gaps 7;

DR PROSITE; PS50041; C_TYPE LECTIN_2; 8.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00213; LIPOCALIN; 1.
DR PROSITE; PS50231; RICIN_B LECTIN; 1.
KW Kinase; Receptor.
SQ SEQUENCE 1479 AA; 166654 MW; C7583EA78E2792D1 CRC64;

Query Match 8.8%; Score 176.5; DB 4; Length 1479;
Best Local Similarity 30.4%; Pred. No. 2.6e-07;
Matches 51; Conservative 25; Mismatches 57; Indels 35; Gaps 7;

QY 28 SQPVCRCGGTQPCYKVIYFHDTSRRLNFEAEKACRRDGGQLVSI ESEDEQKLI EKPIE 87
Db 385 SWQPF-----QGHCYRL-----QAEKRSWQESKKACLRGGDLVSIHSMAELEFITKQIK 434

QY 88 NLLPSDGDGFWIGLRRRREKQSNSTACQDLYAWTDGSI SQFRNWWYVDEPS---CGSEVVCVW 144
Db 435 QEVE---ELWIGL-----NDLKLQMFWSGSLVSFTHWHPFEPNFRDSDL EDCVT 483

QY 145 MYHQPSAPAGIGGPGYMFQWNDRCNMKNFNICKYSDEKPAVPSREAE G 192
Db 484 IW----GPEG-----RWNDSPCNQSLPSICKKAGQLSQGAAEEDHG 520

RESULT 13
Q64449
ID Q64449 PRELIMINARY; PRT; 1479 AA.
AC Q64449;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Lectin lambda.
GN MRC2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96355501; PubMed=8702911;
RA Wu K., Yuan J., Lasky L.A.;
RT "Characterization of a novel member of the macrophage mannose receptor type C lectin family."
RL J. Biol. Chem. 271:21323-21330(1996).
DR EMBL; U56734; AAC52729.1; -.
DR HSSP; P02751; 2FN2.
DR MGD; MGI:107818; Mrc2.
DR InterPro; IPR002353; AntifreezeII.
DR InterPro; IPR001128; Cytochrome P450.
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR000566; Lipocln_cytFABP.
DR InterPro; IPR000772; Ricin_B_lectin.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00059; lectin_c; 8.
DR PRINTS; PR00356; ANTIFREEZEII.
DR PRINTS; PR00013; FNTYPEII.
DR ProDom; PD000995; FN_Type_II; 1.
DR SMART; SM00034; CLECT; 8.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00458; RICIN; 1.
DR PROSITE; PS00086; CYTOCHROME P450; 1.
DR PROSITE; PS00615; C_TYPE LECTIN_1; 3.
DR PROSITE; PS50041; C_TYPE LECTIN_2; 8.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00213; LIPOCALIN; 1.
DR PROSITE; PS50231; RICIN_B LECTIN; 1.
SQ SEQUENCE 1479 AA; 167112 MW; 62D456E10B9B48C1 CRC64;

Query Match 8.5%; Score 170.5; DB 11; Length 1479;
Best Local Similarity 31.4%; Pred. No. 9.4e-07;
Matches 48; Conservative 22; Mismatches 48; Indels 35; Gaps 7;

QY 28 SGQVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIIESEDEQKLIKIEFIE 87
Db 384 SWQPF-----QGHCYRL-----QAEKRSWQESKRACLRGGDLLSIHSMAELEFITKQIK 433
QY 88 NLLPSDGDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNWWYVDEPS---CGSEVCVV 144
Db 434 QEVE---ELWIGL-----NDLKLMNFWSGSLVSFTHWHPFEPNFRDSLEDCVT 482
QY 145 MYHQPSAPAGIGPYMFQWDDRCNMKNFICK 177
Db 483 IW----GPEG-----RWNDSPCNQSLPSICK 504

RESULT 14

Q8C4F8 ID Q8C4F8 PRELIMINARY; PRT; 217 AA.
AC Q8C4F8;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Chondroitin sulfate proteoglycan 3 (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
DR EMBL; AK082298; BAC38458.1; --
FT NON_TER 1
SQ SEQUENCE 217 AA; 25949 MW; 5F1A6A57B05B76D6 CRC64;

Query Match 8.5%; Score 169; DB 11; Length 217;
Best Local Similarity 31.1%; Pred. No. 1e-07;
Matches 42; Conservative 16; Mismatches 47; Indels 30; Gaps 5;
QY 46 YFHDTSRRLNFEEAKEACRRDGGQLVSIIESEDEQKLIKIEFIENLLPSDGDGFWIGLRRREE 105
Db 3 YF---AHRRAWEDAERDCRRRAGHLTSVHSPEEHKFINSP-----GHENSWIGLNDRTV 53
QY 106 QSNSTACQDLYAWTDGSIQFRNWWYVDEPS---CGSEVCVVMYHQPSAPAGIGGPMFQ 162
Db 54 ERD-----FQWTDNTGLQYENWREKQPDNPFAGGEDCVVMVAHESG-----R 95
QY 163 WDDRCNMKNFICK 177
Db 96 WNDVPCNYNLPYVCK 110

RESULT 15

Q8K4Q8 ID Q8K4Q8 PRELIMINARY; PRT; 742 AA.
AC Q8K4Q8;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Collectin placenta 1.
GN COLEC12 OR CL-P1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Ohtani K., Suzuki Y., Eda S., Kawai T., Kase T., Keshi H., Sakai Y.,
RA Fukuoh A., Sakamoto T., Itabe H., Suzutani T., Ogasawara M.,
RA Yoshida I., Wakamiya N.;

RT "cDNA cloning of mouse CL-P1 gene."
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB078434; BAC05523.1; --
DR MGD; MGI:2152907; Colec12.
DR InterPro; IPR002353; AntifreezeII.
DR InterPro; IPR000087; Collagen.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF01391; Collagen; 3.
DR Pfam; PF00059; lectin c; 1.
DR PRINTS; PR00356; ANTIFREEZEII.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C_TYPE_LLECTIN_1; 1.
DR PROSITE; PS50041; C_TYPE_LLECTIN_2; 1.
SQ SEQUENCE 742 AA; 81304 MW; 1537C490E5911C45 CRC64;
Query Match 8.3%; Score 165.5; DB 11; Length 742;
Best Local Similarity 33.8%; Pred. No. 1.1e-06;
Matches 50; Conservative 17; Mismatches 54; Indels 27; Gaps 8;
QY 41 CYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIIESEDEQKLIKIEFIENLLPSDGDGFWIGL 100
Db 618 CY---YF--SLEKEIFEDAKLFCEDKSSHLVFINSRREQQWIKKH-----TVGRESHWIGL 668
QY 101 RRREEKQSNSTACQDLYAWTDGSIQFRNWWYVDEP-SCGSEVCVVMYHQPSAPAGIGGPPY 159
Db 669 TDSEQESE-----WKWLDGSPVDYKNWKAGQPDNWGSG-----HGPGEDCA-GLIY 713
QY 160 MFQWDDRCNMKNFICKYSDEKPAVPS 187
Db 714 AGQWNDFOCDEINNPFCE--KEREAVPS 739
Search completed: December 22, 2003, 16:13:29
Job time : 45.7759 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 22, 2003, 16:05:39 ; Search time 24.0017 Seconds
(without alignments)
2214.797 Million cell updates/sec

Title: US-09-887-855-5
Perfect score: 1115
Sequence: 1 ATGRLLSGQPVCRCGGTQRPC.....EEDAKKTFKESREAAALNLAY 206

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1115	100.0	374	4 Q8TAY8	Q8tay8 homo sapien
2	1115	100.0	374	4 Q96NF3	Q96nf3 homo sapien
3	1111	99.6	374	4 Q96NC5	Q96nc5 homo sapien
4	945.5	84.8	374	11 Q9Z209	Q9z209 cricetulus
5	803.5	72.1	211	11 Q8C351	Q8c351 mus musculus
6	580	52.0	246	11 Q8BVI7	Q8bvi7 mus musculus
7	580	52.0	292	11 Q8BVU2	Q8bvu2 mus musculus
8	185	16.6	1290	13 Q9W6E1	Q9w6e1 gallus gall
9	178.5	16.0	1456	11 Q61830	Q61830 mus musculus
10	177.5	15.9	1348	5 Q25199	Q25199 hydra atten
11	176.5	15.8	1479	4 Q9Y5P9	Q9y5p9 homo sapien
12	176.5	15.8	1479	4 Q9UBG0	Q9ubg0 homo sapien
13	170.5	15.3	1479	11 Q64449	Q64449 mus musculus
14	169	15.2	217	11 Q8C4F8	Q8c4f8 mus musculus
15	165.5	14.8	742	11 Q8K4Q8	Q8k4q8 mus musculus
16	165.5	14.8	742	11 Q8C979	Q8c979 mus musculus

17	161.5	14.5	134	5 Q9XYX3	Q9xyx3 hydra magni
18	159.5	14.3	742	11 Q8VIF6	Q8vif6 mus musculu
19	158.5	14.2	719	6 O62623	O62623 bos taurus
20	158	14.2	142	11 Q8CJ86	Q8cj86 mus musculu
21	158	14.2	142	11 Q8BHK7	Q8bhk7 mus musculu
22	158	14.2	295	11 Q91ZW4	Q91zw4 mus musculu
23	158	14.2	311	11 Q9D8V4	Q9d8v4 mus musculu
24	158	14.2	325	11 Q91ZX0	Q91zx0 mus musculu
25	157	14.1	158	13 Q90WI7	Q90wi7 bungarus fa
26	156	14.0	323	11 Q8CJ91	Q8cj91 mus musculu
27	156	14.0	339	6 Q95244	Q95244 sus scrofa
28	155	13.9	293	11 Q8BGZ0	Q8bgz0 mus musculu
29	155	13.9	323	11 Q8CJ94	Q8cj94 mus musculu
30	155	13.9	323	11 Q8CJ93	Q8cj93 mus musculu
31	155	13.9	323	11 Q8CJ88	Q8cj88 mus musculu
32	155	13.9	1152	13 Q90WM2	Q90wm2 xenopus lae
33	154	13.8	322	11 Q8CJ89	Q8cj89 mus musculu
34	154	13.7	485	6 Q95LG3	Q95lg3 odocoileus
35	152.5	13.7	158	13 Q90WI6	Q90wi6 bungarus mu
36	152	13.6	652	4 Q8IXK1	Q8ixk1 homo sapien
37	151.5	13.6	399	6 Q8HYI2	Q8hyi2 hylobates l
38	150.5	13.5	197	6 Q28008	Q28008 bos taurus
39	150	13.5	158	11 Q8JZX6	Q8jzx6 mus musculu
40	149.5	13.4	459	5 Q22136	Q22136 caenorhabdi
41	149.5	13.4	158	13 Q90WI8	Q90wi8 bungarus fa
42	149	13.4	163	13 Q8AXR8	Q8axr8 anguilla ja
43	148.5	13.3	195	5 Q27340	Q27340 megabalanus
44	148.5	13.3	330	4 Q9NT67	Q9nt67 homo sapien
45	148.5	13.3			

ALIGNMENTS

RESULT 1

Q8TAY8
ID Q8TAY8 PRELIMINARY; PRT; 374 AA.
AC Q8TAY8;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Similar to unnamed protein product.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC025407; AAH25407.1; --
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00059; lectin_c; 1.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS50041; C_TYPE_LLECTIN_2; 1.
SQ SEQUENCE 374 AA; 42312 MW; FC214E6BC9E578D9 CRC64;

Query Match 100.0%; Score 1115; DB 4; Length 374;
Best Local Similarity 100.0%; Pred. No. 4.5e-99;
Matches 206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	ATGRLLSGQPVCRCGGTQRPCYKVIYFHDTSRRLLNFEEAKEACRRDGGQLVSI	60
Db	22	ATGRLLSGQPVCRCGGTQRPCYKVIYFHDTSRRLLNFEEAKEACRRDGGQLVSI	81
Qy	61	TEKFIENLLPSDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNWWVDEPSCGSEV	120
Db	82	TEKFIENLLPSDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNWWVDEPSCGSEV	141
Qy	121	CVVMYHQPAPAGIGGYPYMFQWDDRCNMKNFICKYSDKPAVPSREAEGETELTPV	180
Db	142	CVVMYHQPAPAGIGGYPYMFQWDDRCNMKNFICKYSDKPAVPSREAEGETELTPV	201

QY 181 LPEETQEDAKKTFKESREAAALNLAY 206
Db 202 LPEETQEDAKKTFKESREAAALNLAY 227

RESULT 2
Q96NF3 PRELIMINARY; PRT; 374 AA.
AC Q96NF3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein FLJ30977.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Ishibashi T., Kanehori K., Yosida M., Watanabe S., Ishida S., Ono Y.,
RA Hotuta T., Hiraoka S., Murakawa K., Takiguchi S., Kusano J.,
RA Watanabe M., Fujimori K., Tanai H., Ishida M., Yamashita H., Chiba Y.,
RA Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S.,
RA Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
RA Kimura K., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
RA Wagatsuma M., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
RA Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK055539; BAB70946.1; -.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00059; lectin_c; 1.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW Hypothetical protein.
SQ SEQUENCE 374 AA; 42280 MW; 8AE64E6BC9E56DCD CRC64;

Query Match 100.0%; Score 1115; DB 4; Length 374;
Best Local Similarity 100.0%; Pred. No. 4.5e-99;
Matches 206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLLNFEEAKEACRRDGGQLVSI ESEDEQKL 60
Db 22 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLLNFEEAKEACRRDGGQLVSI ESEDEQKL 81

QY 61 IEKFIENLLPSDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNWWYVDEPSCGSEV 120
Db 82 IEKFIENLLPSDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNWWYVDEPSCGSEV 141

QY 121 CVVMYHQPSAPAGIGGYPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGETELTTPV 180
Db 142 CVVMYHQPSAPAGIGGYPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGETELTTPV 201

QY 181 LPEETQEDAKKTFKESREAAALNLAY 206
Db 202 LPEETQEDAKKTFKESREAAALNLAY 227

RESULT 3
Q96NC5 PRELIMINARY; PRT; 374 AA.
AC Q96NC5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein FLJ31092.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.

RA Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,
RA Fukuzumi Y., Fujimori Y., Komiyama M., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,
RA Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A.,
RA Kawakami B., Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K.,
RA Isogai T.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK055654; BAB70978.1; -.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00059; lectin_c; 1.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW Hypothetical protein.
SQ SEQUENCE 374 AA; 42310 MW; CBF74E676E23BA16 CRC64;

Query Match 99.6%; Score 1111; DB 4; Length 374;
Best Local Similarity 99.5%; Pred. No. 1.1e-98;
Matches 205; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLLNFEEAKEACRRDGGQLVSI ESEDEQKL 60
Db 22 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLLNFEEAKEACRRDGGQLVSI ESEDEQKL 81

QY 61 IEKFIENLLPSDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNWWYVDEPSCGSEV 120
Db 82 IEKFIENLLPSDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNWWYVDEPSCGSEV 141

QY 121 CVVMYHQPSAPAGIGGYPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGETELTTPV 180
Db 142 CVVMYHQPSAPAGIGGYPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGETELTTPV 201

QY 181 LPEETQEDAKKTFKESREAAALNLAY 206
Db 202 LPEETQEDAKKTFKESREAAALNLAY 227

RESULT 4
Q9Z209 PRELIMINARY; PRT; 374 AA.
AC Q9Z209;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Layilin.
OS Cricetulus griseus (Chinese hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Cricetulus.
OX NCBI_TaxID=10029;
RN [1]
RP SEQUENCE FROM N.A.
RA Borowsky M.L., Hynes R.O.;
RT "Layilin, a novel talin-binding transmembrane protein homologous with
RT C-type lectins, is localized in membrane ruffles.";
RL J. Cell Biol. 143:0-0(1998).
DR EMBL; AF093673; AAC68695.1; -.
DR HSSP; P22897; iEGG.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00059; lectin_c; 1.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
SQ SEQUENCE 374 AA; 42435 MW; 298A8BA24FB04E1C CRC64;

Query Match 84.8%; Score 945.5; DB 11; Length 374;
Best Local Similarity 85.9%; Pred. No. 9.4e-83;
Matches 176; Conservative 9; Mismatches 19; Indels 1; Gaps 1;

QY 3 GRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLLNFEEAKEACRRDGGQLVSI ESEDEQKLIE 62
Db 24 GRLLSGQLVCRGGTRRPCYKVIYFHDAFQRLNFEEAKEACRRDGGQLVSI ETEDEQRLIE 83

Db 955 CFKIFGFANEKKSQDARQACKGL 979

RESULT 10

Q25199 ID Q25199 PRELIMINARY; PRT; 1348 AA.

AC Q25199

DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Tyrosine kinase receptor.

OS Hydra attenuata (Hydra) (Hydra vulgaris).

OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Anthomedusae;

OC Hydridae; Hydra.

OX NCBI_TaxID=6087;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Irvine;

RX MEDLINE=20209407; PubMed=10744720;

RA Reidling J.C., Miller M.A., Steele R.E.;

RT "Sweet Tooth, a Novel Receptor Protein-tyrosine Kinase with C-type

RT Lectin-like Extracellular Domains.";

RL J. Biol. Chem. 275:10323-10330(2000).

DR EMBL; L22612; AAA29218.2; -.

DR HSSP; P11362; IFGK.

DR InterPro; IPR001304; Lectin_C.

DR InterPro; IPR003990; Pancreatias_ac.

DR InterPro; IPR000719; Prot_kinase.

DR InterPro; IPR001245; Tyr_kinase.

DR Pfam; PF00059; lectin_c; 4.

DR Pfam; PF00069; pkinase; 1.

DR PRINTS; PR01504; PNCREATITSAP.

DR ProDom; PD000001; Prot_kinase; 2.

DR SMART; SM00034; CLECT; 4.

DR SMART; SM00219; TyrKC; 1.

DR PROSITE; PS00615; C_TYPE_LLECTIN_1; 2.

DR PROSITE; PS50041; C_TYPE_LLECTIN_2; 4.

DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.

DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.

DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.

KW ATP-binding; Kinase; Receptor; Transferase; Tyrosine-protein kinase.

SQ SEQUENCE 1348 AA; 156916 MW; 77D2122093227FEF CRC64;

Query Match 15.9%; Score 177.5; DB 5; Length 1348;

Best Local Similarity 28.0%; Pred. No. 3.9e-08;

Matches 47; Conservative 32; Mismatches 46; Indels 43; Gaps 8;

Qy 6 LSGQPVC--RGGTQPC-----YKVIYFHDTSRRLLNFEEAKEACRRDGGQLVSIESTE 57

Db 412 LSHRFICKVKRATNEYCAEGWTSYRIYCYFIYSIEFDWFKSFSCQIGNGLLSIENQEE 471

Qy 58 QKLIKFIENLLPSDGD-FWIGLRR-----REEKQSNSTACQDLYAWTDGSIQFRNWWY 110

Db 472 ----NRFIENDLIKNDKYYWIGLNKIWNLYLKNKR-----FEWSDNTYTQFFNWI 518

Qy 111 VDEP--SCGSEVCVVMYHQPSAPAGIGGPMFQWNDRCNMKNPFICK 156

Db 519 TNQPDNNGIESCVMNYN-----GWSDKCKVLNGFICK 553

RESULT 11

Q9Y5P9 ID Q9Y5P9 PRELIMINARY; PRT; 1479 AA.

AC Q9Y5P9

DT 01-NOV-1999 (TrEMBLrel. 12, Created)

DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Endocytic receptor Endol80.

GN ENDO180.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Behrendt N., Jensen O.N., Engelholm L.H., Mortz E., Mann M., Dano K.;

RT "A urokinase receptor-associated protein with specific collagen-

RT binding properties.";

RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Brain;

RX MEDLINE=98403880; PubMed=9734811;

RA Ishikawa K., Nagase T., Suyama M., Miyajima N., Tanaka A., Kotani H.;

RA Nomura N., Ohara O.;

RT "Prediction of the coding sequences of unidentified human genes. X.

RT The complete sequences of 100 new cDNA clones from brain which can

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=20148849; PubMed=10683150;

RA Sheikh H., Yarwood H., Ashworth A., Isacke C.;

RT "Endol80, an endocytic recycling glycoprotein related to the

RT macrophage mannose receptor is expressed on fibroblasts, endothelial

RT cells and macrophages and functions as a lectin receptor.";

RL J. Cell Sci. 113:1021-1032(2000).

DR EMBL; AF134838; AAD30280.1; -.

DR HSSP; P02751; 2FN2.

DR InterPro; IPR000562; FN_Type_II.

DR InterPro; IPR001304; Lectin_C.

DR InterPro; IPR000566; Lipocln_cytFABP.

DR InterPro; IPR000772; Ricin_B_lectin.

DR Pfam; PF00040; fn2; 1.

DR Pfam; PF00059; lectin_c; 8.

DR PRINTS; PR00013; FNTYPEII.

DR ProDom; PD000995; FN_Type_II; 1.

DR SMART; SM00034; CLECT; 8.

DR SMART; SM00059; FN2; 1.

DR SMART; SM00458; RICIN; 1.

DR PROSITE; PS00615; C_TYPE_LLECTIN_1; 3.

DR PROSITE; PS50041; C_TYPE_LLECTIN_2; 8.

DR PROSITE; PS00023; FIBRONECTIN_2; 1.

DR PROSITE; PS00213; LIPOCALIN; 1.

DR PROSITE; PS50231; RICIN_B_LLECTIN; 1.

KW Receptor.

SQ SEQUENCE 1479 AA; 166669 MW; 9F4BAF355F036FCE CRC64;

Query Match 15.8%; Score 176.5; DB 4; Length 1479;

Best Local Similarity 30.4%; Pred. No. 5.4e-08;

Matches 51; Conservative 25; Mismatches 57; Indels 35; Gaps 7;

Qy 7 SGQPVCRGGTQPCYKVIYFHDTSRRLLNFEEAKEACRRDGGQLVSIESTEQLIEKFIE 66

Db 385 SWQPF-----QGHCYRL-----QAEKRSWQESKKACLRGGGDLVSIHMAELEFITKQIK 434

Qy 67 NLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNWWYVDEPS---CGSEVCVW 123

Db 435 QEVE---ELWIGL-----NDLKLQMNFEWSDGSLVSFTHWHPFPNPNFRDSDLEDCVT 483

Qy 124 MYHQPSAPAGIGGPMFQWNDRCNMKNPFICKYSDEKPAVPSREAEG 171

Db 484 IW---GPEG-----RWNDSPCNQSLPSICKKAGQLSQGAEEHDHG 520

RESULT 12

Q9UBG0 ID Q9UBG0 PRELIMINARY; PRT; 1479 AA.

AC Q9UBG0

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Urokinase receptor-associated protein UPARAP.

GN KIAA0709.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Behrendt N., Jensen O.N., Engelholm L.H., Mortz E., Mann M., Dano K.;

RT "A urokinase receptor-associated protein with specific collagen-

RT binding properties.";

RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Brain;

RX MEDLINE=98403880; PubMed=9734811;

RA Ishikawa K., Nagase T., Suyama M., Miyajima N., Tanaka A., Kotani H.;

RA Nomura N., Ohara O.;

RT "Prediction of the coding sequences of unidentified human genes. X.

RT The complete sequences of 100 new cDNA clones from brain which can

RT code for large proteins in vitro.;
RL DNA Res. 5:169-176(1998).
DR EMBL; AF107292; AAF14192.1; -.
DR EMBL; AB014609; BAA31684.1; -.
DR HSSP; P02751; 2FN2.
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR000566; Lipocln_cytFABP.
DR InterPro; IPR000772; Ricin_B_lectin.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00059; lectin_c; 8.
DR PRINTS; PRO0013; FNTYPEII.
DR ProDom; PD000995; FN_Type_II; 1.
DR SMART; SM00034; CLECT; 8.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00458; RICIN; 1.
DR PROSITE; PS00615; C_TYPE_LLECTIN_1; 3.
DR PROSITE; PS00041; C_TYPE_LLECTIN_2; 8.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00213; LIPOCALIN; 1.
DR PROSITE; PS0231; RICIN_B_LLECTIN; 1.
KW Kinase; Receptor.
SQ SEQUENCE 1479 AA; 166654 MW; C7583EA78E2792D1 CRC64;

Query Match 15.8%; Score 176.5; DB 4; Length 1479;
Best Local Similarity 30.4%; Pred. No. 5.4e-08;
Matches 51; Conservative 25; Mismatches 57; Indels 35; Gaps 7;

QY 7 SQQPVCRGGTQPCYKVIYFHDTSRRLLNFEEAKEACRRDGGQLVSIIESEDEQKLEKPIE 66
Db 385 SQQPF-----QGHCYRL-----QAEKRSWQESKKACLRGGGDLVSIHSMAELEFITKQIK 434

QY 67 NLLPSDGFDFWIGLRRRREKQSNSTACQDLYAWTDGSIQFRNWWYVDEPS---CGSEVVCV 123
Db 435 QEVE---ELWIGL-----NDLKLQMNFEWSGSLVSFTHWHHPFEPNFRDLSLEDCVT 483

QY 124 MYHQPSAPAGIGGYPYMFQWDDRCNMKNFICKYSDEKPAVPSREAE 171
Db 484 IW----GPEG-----RWNDSPCQNSLPSICKKAGQLSQGAEEEDHG 520

RESULT 13
Q64449 PRELIMINARY; PRT; 1479 AA.
AC Q64449;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Lectin lambda.
GN MRC2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96355501; PubMed=8702911;
RA Wu K., Yuan J., Lasky L.A.;
RT "Characterization of a novel member of the macrophage mannose receptor
RT type C lectin family.";
RL J. Biol. Chem. 271:21323-21330(1996).
DR EMBL; U56734; AAC52729.1; -.
DR HSSP; P02751; 2FN2.
DR MGD; MGI:107818; Mrc2.
DR InterPro; IPR002353; AntifreezeII.
DR InterPro; IPR001128; Cytochrome P450.
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR000566; Lipocln_cytFABP.
DR InterPro; IPR000772; Ricin_B_lectin.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00059; lectin_c; 8.
DR PRINTS; PR00356; ANTIFREEZEII.

DR PRINTS; PR00013; FNTYPEII.
DR ProDom; PD000995; FN_Type_II; 1.
DR SMART; SM00034; CLECT; 8.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00458; RICIN; 1.
DR PROSITE; PS00086; CYTOCHROME P450; 1.
DR PROSITE; PS00615; C_TYPE_LLECTIN_1; 3.
DR PROSITE; PS00041; C_TYPE_LLECTIN_2; 8.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00213; LIPOCALIN; 1.
DR PROSITE; PS0231; RICIN_B_LLECTIN; 1.
SQ SEQUENCE 1479 AA; 167112 MW; 62D456E108B9B48C1 CRC64;

Query Match 15.3%; Score 170.5; DB 11; Length 1479;
Best Local Similarity 31.4%; Pred. No. 2.1e-07;
Matches 48; Conservative 22; Mismatches 48; Indels 35; Gaps 7;

QY 7 SQQPVCRGGTQPCYKVIYFHDTSRRLLNFEEAKEACRRDGGQLVSIIESEDEQKLEKPIE 66
Db 384 SQQPF-----QGHCYRL-----QAEKRSWQESKKACLRGGGDLVSIHSMAELEFITKQIK 433

QY 67 NLLPSDGFDFWIGLRRRREKQSNSTACQDLYAWTDGSIQFRNWWYVDEPS---CGSEVVCV 123
Db 434 QEVE---ELWIGL-----NDLKLQMNFEWSGSLVSFTHWHHPFEPNFRDLSLEDCVT 482

QY 124 MYHQPSAPAGIGGYPYMFQWDDRCNMKNFICK 156
Db 483 IW----GPEG-----RWNDSPCQNSLPSICK 504

RESULT 14
Q8C4F8 PRELIMINARY; PRT; 217 AA.
AC Q8C4F8;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Chondroitin sulfate proteoglycan 3 (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK082298; BAC38458.1; -.
FT NON TER 1
SQ SEQUENCE 217 AA; 25949 MW; 5F1A6A57B05B76D6 CRC64;

Query Match 15.2%; Score 169; DB 11; Length 217;
Best Local Similarity 31.1%; Pred. No. 2.9e-08;
Matches 42; Conservative 16; Mismatches 47; Indels 30; Gaps 5;

QY 25 YFHDTSRRLLNFEEAKEACRRDGGQLVSIIESEDEQKLEKFIENLLPSDGFDFWIGLRRREE 84
Db 3 YF---AHRRAWEDAERDCRRRAGHLTSVHSPEEHKFINSF-----GHENSWIGLNDRTV 53

QY 85 KQSNSTACQDLYAWTDGSIQFRNWWYVDEPS---CGSEVVCVVMYHQPSAPAGIGGYPYMFQ 141
Db 54 ERD-----FQWTDNTGLQYENWREKQPDNFFAGGEDCVVMVAHESG-----R 95

QY 142 WNDRCNMKNFICK 156
Db 96 WNDVPCNYNLPYVCK 110

RESULT 15

Q8K4Q8			
ID	Q8K4Q8	PRELIMINARY;	PRT; 742 AA.
AC	Q8K4Q8;		
DT	01-OCT-2002 (Tremblrel. 22, Created)		
DT	01-OCT-2002 (Tremblrel. 22, Last sequence update)		
DT	01-MAR-2003 (Tremblrel. 23, Last annotation update)		
DE	Collectin placenta 1.		
GN	COLEC12 OR CL-P1.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Ohtani K., Suzuki Y., Eda S., Kawai T., Kase T., Keshi H., Sakai Y.,		
RA	Fukuoh A., Sakamoto T., Itabe H., Suzutani T., Ogasawara M.,		
RA	Yoshida I., Wakamiya N.;		
RT	"cDNA cloning of mouse CL-P1 gene.";		
RL	Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AB078434; BAC05523.1; -		
DR	MGD; MGI:2152907; Colec12.		
DR	InterPro; IPR002353; AntifreezeII.		
DR	InterPro; IPR000087; Collagen.		
DR	InterPro; IPR001304; Lectin_C.		
DR	Pfam; PF01391; Collagen; 3.		
DR	Pfam; PF00059; lectin c; 1.		
DR	PRINTS; PR00356; ANTIFREEZEII.		
DR	SMART; SM00034; CLECT; 1.		
DR	PROSITE; PS00615; C_TYPE_LLECTIN_1; 1.		
DR	PROSITE; PS50041; C_TYPE_LLECTIN_2; 1.		
SO	SEQUENCE 742 AA; 81304 MW; 1537C490E5911C45 CRC64;		
Query Match 14.8%; Score 165.5; DB 11; Length 742;			
Best Local Similarity 33.8%; Pred. No. 2.7e-07;			
Matches 50; Conservative 17; Mismatches 54; Indels 27; Gaps 8;			
QY	20	CYKVIYFHDTSRRLNFEFEAKEACRRDGGQLVSISEDEQKLIKFIENLLPSDGDWIGL 79	
Db	618	CY---YF--SLEKEIFEAKLFCEDKSSHVLVFINSREEQQWIKKH----TVGRESHWIGL 668	
QY	80	RRREEKQSNSTACQDLYAWTDGSIQFRNWWYVDEP-SCGSEVCVVMYHQPSAPAGIGGPY 138	
Db	669	TDSEQESE-----WKWLDGSPVDYKNWKAGQPDNWSG-----HGPGEDCA-GLIY 713	
QY	139	MFQWDDRCNMKNFICKYSDEKPAVPS 166	
Db	714	AGQWDFQCDEINNFI--KEREAVPS 739	

Search completed: December 22, 2003, 16:13:31
Job time : 24.2017 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 22, 2003, 16:04:13 ; Search time 52.9807 Seconds
(without alignments)
1120.478 Million cell updates/sec

Title: US-09-887-855-2

Perfect score: 2000

Sequence: 1 MRPQTALQAVLLAVLLVGLR.....PDQMGRSKESGWENEIYGY 374

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- A Geneseq 19Jun03:*
- 1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
 - 2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
 - 3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
 - 4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
 - 5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
 - 6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
 - 7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
 - 8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
 - 9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
 - 10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
 - 11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
 - 12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
 - 13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
 - 14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
 - 15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
 - 16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
 - 17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
 - 18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
 - 19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
 - 20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
 - 21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
 - 22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
 - 23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
 - 24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2000	100.0	374	21	AA93948
2	2000	100.0	374	22	AAE03651
3	2000	100.0	374	23	ABB90203
4	1990	99.5	374	21	AA91490
5	1986	99.3	382	20	AA13367
6	1986	99.3	382	22	AAU29033
7	1986	99.3	382	22	AA80235
8	1986	99.3	382	24	ABU69645
9	1986	99.3	382	24	ABU71121

Appended - checked

10	1986	99.3	382	24	ABU71468	Human PRO polypept
11	1986	99.3	382	24	ABU71914	Human secreted/tra
12	1986	99.3	382	24	ABU65578	Human secreted/tra
13	1986	99.3	382	24	ABU65911	Novel human secret
14	1986	99.3	382	24	ABU67368	Human secreted pro
15	1986	99.3	382	24	ABU67415	Human secreted/tra
16	1986	99.3	382	24	ABU64522	Human secreted/tra
17	1986	99.3	382	24	ABU65273	Human PRO polypept
18	1986	99.3	382	24	ABU58409	Human PRO polypept
19	1986	99.3	382	24	ABU55945	Human secreted/tra
20	1986	99.3	382	24	ABU56940	Human PRO polypept
21	1986	99.3	382	24	ABU54370	Human secreted/tra
22	1986	99.3	382	24	ABU10519	Human secreted/tra
23	1982	99.1	387	22	AA25796	Human protein sequ
24	1979	99.0	374	23	ABG6680	Human novel polype
25	645	32.2	273	21	AA18913	A novel polypeptid
26	645	32.2	273	22	AAU12441	Human PRO1890 poly
27	645	32.2	273	22	AA873309	Human C-type lecti
28	645	32.2	273	22	AA887609	Human PRO1890. Ho
29	645	32.2	273	23	ABG95934	Human secreted/tra
30	645	32.2	273	23	ABB95554	Human angiogenesis
31	645	32.2	273	23	ABB84948	Human PRO1890 prot
32	645	32.2	273	24	ABU69084	Human PRO polypept
33	645	32.2	273	24	ABU69107	Human PRO polypept
34	645	32.2	273	24	ABU71589	Human secreted pol
35	645	32.2	273	24	ABU72035	Novel human secret
36	645	32.2	273	24	ABU72192	Human PRO polypept
37	645	32.2	273	24	ABU66839	Human PRO polypept
38	645	32.2	273	24	ABU67115	Human secreted/tra
39	645	32.2	273	24	ABU59920	Novel secreted and
40	606	30.3	274	23	ABP69211	Human polypeptide
41	576	28.8	232	22	AA894192	Human protein sequ
42	336	16.8	81	21	AA91643	Human secreted pro
43	336	16.8	82	21	AA91557	Human secreted pro
44	336	16.8	115	21	AA91642	Human secreted pro
45	275	13.8	102	22	ABG51278	Human liver peptid

ALIGNMENTS

RESULT 1
AA93948

ID AA93948 standard; Protein; 374 AA.

XX

AC AA93948;

XX 03-OCT-2000 (first entry)

DT Amino acid sequence of a lectin ss3939 polypeptide.

XX Human; lectin ss3939; chromosome 11; gene therapy.

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT Peptide 1..21

FT Domain /note= "signal peptide"

FT Domain 22..227

FT Domain /note= "extracellular coding region"

FT Domain 228..248

FT Domain /note= "predicted transmembrane domain"

FT Domain 249..374

FT Domain /note= "predicted cytoplasmic or intracellular domain"

XX WO200039296-A1.

PN 06-JUL-2000.

XX 22-DEC-1999; 99WO-US30523.

XX 23-DEC-1998; 98US-0113820.

CC glomerulonephritis, multiple sclerosis, Grave's disease, osteoarthritis,
CC osteoporosis, psoriasis, rheumatoid arthritis, ulcerative colitis,
CC bacterial, fungal, parasitic, protozoal and helminthic infections, and
CC cell proliferative disorders such as actinic keratosis, arteriosclerosis
CC and cancer including breast, bladder, bone marrow, brain and uterus
CC cancer, leukaemia, adenocarcinoma, lymphoma, melanoma and myeloma.
XX
SQ Sequence 374 AA;
Query Match 100.0%; Score 2000; DB 22; Length 374;
Best Local Similarity 100.0%; Pred. No. 3.1e-184;
Matches 374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQPCVKYVFHDTSRRLNFEEAK 60
Db 1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQPCVKYVFHDTSRRLNFEEAK 60
Qy 61 EACRRDGGQLVSI ESEDEQKLI EKFIENLLPSDGDGFWIGLRRRREKQSNSTACQDLYAWT 120
Db 61 EACRRDGGQLVSI ESEDEQKLI EKFIENLLPSDGDGFWIGLRRRREKQSNSTACQDLYAWT 120
Qy 121 DGSISQFRNWWYVDEPSCGSEVVCVMYHQPSPAPAGIGGYPYMFQWDDRCNMKNFICKYSD 180
Db 121 DGSISQFRNWWYVDEPSCGSEVVCVMYHQPSPAPAGIGGYPYMFQWDDRCNMKNFICKYSD 180
Qy 181 EKPAVPSREAEGETELTTPVLPEETQEDAKKTFKESREAAALNLAYILIPSIPLLLLLV 240
Db 181 EKPAVPSREAEGETELTTPVLPEETQEDAKKTFKESREAAALNLAYILIPSIPLLLLLV 240
Qy 241 VTTVVCWWVICRKRKREQDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300
Db 241 VTTVVCWWVICRKRKREQDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300
Qy 301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMG 360
Db 301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMG 360
Qy 361 SKESGWVENEIYGY 374
Db 361 SKESGWVENEIYGY 374
RESULT 3
ABB90203
ID ABB90203 standard; Protein; 374 AA.
XX
AC ABB90203;
XX
DT 24-MAY-2002 (first entry)
XX
DE Human polypeptide SEQ ID NO 2579.
XX
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
KW vulnery; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein.
OS Homo sapiens.
XX
PN WO200190304-A2.
XX
PD 29-NOV-2001.
XX
PF 18-MAY-2001; 2001WO-US16450.
XX
PR 19-MAY-2000; 2000US-205515P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Birse CE, Rosen CA;
XX
DR WPI; 2002-122018/16.

DR N-PSDB; ABL90612.
XX
PT Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
PT prevention of neural, immune system, muscular, reproductive,
PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
PT disorders.
XX
PS Claim 11; SEQ ID NO 2579; 2081pp + Sequence Listing; English.
XX
CC The invention relates to novel genes (ABL89449-ABL90853) and proteins
CC (ABB9040-ABB9044) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 374 AA;
Query Match 100.0%; Score 2000; DB 23; Length 374;
Best Local Similarity 100.0%; Pred. No. 3.1e-184;
Matches 374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQPCVKYVFHDTSRRLNFEEAK 60
Db 1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQPVCRGGTQPCVKYVFHDTSRRLNFEEAK 60
Qy 61 EACRRDGGQLVSI ESEDEQKLI EKFIENLLPSDGDGFWIGLRRRREKQSNSTACQDLYAWT 120
Db 61 EACRRDGGQLVSI ESEDEQKLI EKFIENLLPSDGDGFWIGLRRRREKQSNSTACQDLYAWT 120
Qy 121 DGSISQFRNWWYVDEPSCGSEVVCVMYHQPSPAPAGIGGYPYMFQWDDRCNMKNFICKYSD 180
Db 121 DGSISQFRNWWYVDEPSCGSEVVCVMYHQPSPAPAGIGGYPYMFQWDDRCNMKNFICKYSD 180
Qy 181 EKPAVPSREAEGETELTTPVLPEETQEDAKKTFKESREAAALNLAYILIPSIPLLLLLV 240
Db 181 EKPAVPSREAEGETELTTPVLPEETQEDAKKTFKESREAAALNLAYILIPSIPLLLLLV 240
Qy 241 VTTVVCWWVICRKRKREQDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300
Db 241 VTTVVCWWVICRKRKREQDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300
Qy 301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMG 360
Db 301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMG 360
Qy 361 SKESGWVENEIYGY 374
Db 361 SKESGWVENEIYGY 374
RESULT 4
AA91490
ID AA91490 standard; Protein; 374 AA.
XX
AC AA91490;
XX
DT 29-JUN-2000 (first entry)
XX
DE Human secreted protein sequence encoded by gene 40 SEQ ID NO:163.
XX

KW Human; secreted protein; diagnosis; cytostatic; immunosuppressive;
KW antiHIV; antiinflammatory; nootropic; neuroprotective; antiallergic;
KW osteopathic; antiarthritic; antibacterial; antidiabetic; antiasthma;
KW antipsoriatic; cardiant; gene therapy; cancer; neurological disorder;
KW immune disease; inflammation; blood disorder; tumour.
XX
OS Homo sapiens.
XX
XX WO200006698-A1.
PN
XX
XX 10-FEB-2000.
PD
XX
XX 29-JUL-1999; 99WO-US17130.
PF
XX
XX 30-JUL-1998; 98US-0094657.
PR 05-AUG-1998; 98US-0095486.
PR 06-AUG-1998; 98US-0095454.
PR 06-AUG-1998; 98US-0095455.
PR 12-AUG-1998; 98US-0096319.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX Komatsoulis GA, Rosen CA, Ruben SM, Duan R, Moore PA, Shi Y;
PI Lafleur D, Wei Y, Ni J, Florence KA, Young PE, Brewer LA;
PI Soppet DR, Endress GA, Ebner R, Olsen HS, Mucenski M;
XX
DR WPI; 2000-195282/17.
DR N-PSDB; AAA26385.
XX
XX New isolated human genes and the secreted polypeptides they encode,
PT useful for diagnosis and treatment of e.g. cancers, neurological
PT disorders, immune diseases, inflammation or blood disorders -
XX
XX Claim 11; Page 483-484; 634pp; English.
XX
XX The polynucleotide sequences given in AAA26346 to AAA26458 encode the
CC human secreted proteins given in AAY91451 to AAY91691. The human secreted
CC proteins can have activities based on the tissues and cells they are
CC expressed in. Examples of the activities are: cytostatic;
CC immunosuppressive; antiHIV; antiinflammatory; nootropic; neuroprotective;
CC antiallergic; osteopathic; antiarthritic; antibacterial; antidiabetic;
CC antiasthma; antipsoriatic; and cardiant. The polynucleotides and their
CC corresponding secreted proteins are useful for preventing, treating or
CC ameliorating medical conditions, e.g. by protein or gene therapy. Also
CC pathological conditions can be diagnosed by determining the amount of the
CC proteins in a sample or by determining the presence of mutations in the
CC polynucleotides. Specific uses are described for each of the
CC polynucleotides, based on which tissues they are most highly expressed
CC in, and include developing products for the diagnosis or treatment of
CC cancer, tumours, neurodegenerative disorders, developmental abnormalities
CC and foetal deficiencies, hepatic and renal disease, inflammation,
CC autoimmune diseases, hepatic and renal disease, inflammation,
CC allergies, Alzheimer's and behavioural disorders, schizophrenia,
CC osteoporosis, arthritis, infections, AIDS, spinal cord injuries,
CC transplant rejection, diabetes, asthma, sepsis, acne, psoriasis,
CC cardiovascular disorders, reproductive disorders, gastrointestinal
CC disorders, respiratory disorders and metabolic disorders. The proteins
CC or polynucleotides can also be used as food additives or preservatives.
CC The proteins are also useful for identifying their binding partners.
CC AAA26337 to AAA26345 and AAY91450 are sequences used in the
CC exemplification of the present invention.
XX
SQ Sequence 374 AA;

Query Match 99.5%; Score 1990; DB 21; Length 374;
Best Local Similarity 99.5%; Pred. No. 2.9e-183;
Matches 372; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQVCRGGTQPCYKVIYFHDTSRRLNFEAK 60
Db 1 MRPGTALQAVLLAVLLVGLRAATGRLLSGQVCRGGTQPCYKVIYFHDTSRRLNFEAK 60
QY 61 EACRRDGGQLVSI ESEDEQKLI EKFIENLLPSD GDFWIGLRRRREKQSNSTACQDLYAWT 120

Db 61 EACRRDGGQLVSI ESEDEQKLI EKFIENLLPSD GDFWIGLRRRREKQSNSTACQDLYAWT 120
QY 121 DGSISQFRNYYVDEPSCGSEVCVVMYHQPSAPAGIGGYPYMFQWDDRCNMKNFICKYSD 180
Db 121 DGSISQFRNYYVDEPSCGSEVCVVMYHQPSAPAGIGGYPYMFQWDDRCNMKNFICKYSD 180
QY 181 EKPAVPSREAEGBETELTPVLPEETQEEDAKKTFKESREAAALNLAYILIPSIPULLLLLV 240
Db 181 EKPAVPSREAEGBETELTPVLPEETQEEDAKKTFKESREAAALNLAYILIPSIPULLLLLV 240
QY 241 VTTVVCWVWICRKREKREQDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300
Db 241 VTTVVCWVWICRKREKREQDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRP 300
QY 301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMR 360
Db 301 DLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMR 360
QY 361 SKESGWVENEIYGY 374
Db 361 SKESGWVENEIYGY 374
RESULT 5
AAY13367
ID AAY13367 standard; Protein; 382 AA.
XX
AC AAY13367;
XX
DT 25-JUN-1999 (first entry)
XX
DE Amino acid sequence of protein PRO234.
XX
KW Secreted protein; transmembrane protein; human; enterocolitis;
KW Zollinger-Ellison syndrome; gastrointestinal ulceration;
KW congenital microvillus atrophy; skin disease; cell growth;
KW abnormal keratinocyte differentiation; psoriasis; epithelial cancer;
KW Parkinson's disease; Alzheimer's disease; ALS; neuropathy;
KW fibromodulin; dermal scarring; Usher Syndrome; Atrophia areata;
KW anti-thrombotic; wound healing; tissue repair.
OS Homo sapiens.
XX
XX WO9914328-A2.
PN
PD
XX 25-MAR-1999.
PF 16-SEP-1998; 98WO-US19330.
XX
PR 25-NOV-1997; 97US-0066840.
PR 17-SEP-1997; 97US-0059113.
PR 17-SEP-1997; 97US-0059115.
PR 17-SEP-1997; 97US-0059117.
PR 17-SEP-1997; 97US-0059119.
PR 17-SEP-1997; 97US-0059121.
PR 17-SEP-1997; 97US-0059122.
PR 17-SEP-1997; 97US-0059184.
PR 18-SEP-1997; 97US-0059263.
PR 18-SEP-1997; 97US-0059266.
PR 15-OCT-1997; 97US-0062125.
PR 17-OCT-1997; 97US-0062285.
PR 17-OCT-1997; 97US-0062287.
PR 21-OCT-1997; 97US-0063486.
PR 24-OCT-1997; 97US-0062814.
PR 24-OCT-1997; 97US-0062816.
PR 24-OCT-1997; 97US-0063045.
PR 24-OCT-1997; 97US-0063120.
PR 24-OCT-1997; 97US-0063121.
PR 24-OCT-1997; 97US-0063127.
PR 24-OCT-1997; 97US-0063128.
PR 27-OCT-1997; 97US-0063329.
PR 27-OCT-1997; 97US-0063327.

PR 25-APR-2000; 2000US-199654P.
PR 03-MAY-2000; 2000US-201516P.
PR 17-MAY-2000; 2000WO-US13705.
PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.
PR 02-JUN-2000; 2000WO-US15264.
PR 05-JUN-2000; 2000US-209832P.
PR 28-JUL-2000; 2000WO-US20710.
PR 22-AUG-2000; 2000US-0644848.
PR 24-AUG-2000; 2000WO-US23328.
PR 08-NOV-2000; 2000WO-US30952.
PR 01-DEC-2000; 2000WO-US32678.
PR 20-DEC-2000; 2000WO-US34956.
XX
XX (GETH) GENENTECH INC.
XX
PI Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;
PI Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;
XX
DR WPI; 2001-602746/68.
DR N-PSDB; AAS45934.
XX
PT Novel nucleic acids encoding PRO polypeptides, used to diagnose the
PT presence of tumours, such as prostate and breast tumours, in mammals and
PT to screen for modulators of the compounds -
XX
PS Claim 11; Fig 20; 774pp; English.
XX
CC Sequences AAU29024-AAU29328 represent PRO polypeptides of the invention.
CC The PRO polypeptides and their associated nucleic acids can be used to
CC detect the presence of a tumour in a mammal by comparing the level of
CC expression of a PRO polypeptide in a test sample of cells from the animal
CC and a control sample of normal cells, whereby a higher level of
CC expression in the test sample indicates the presence of a tumour in the
CC mammal. Mammals include dogs, cats, cattle, horses, sheep, pigs, goats
CC and rabbits but are preferably human. The polypeptides can be used to
CC stimulate tumour necrosis factor (TNF) alpha release from human blood,
CC when contacted with it. A specific polypeptide can be used to stimulate
CC the proliferation or differentiation of chondrocyte cells. The PRO
CC proteins can be used to determine the presence of tumours and also
CC susceptibility to tumour development, particularly adrenal, lung, colon,
CC breast, prostate, cervical, or liver tumours, in mammalian
CC subjects. The oligonucleotide probes specific for the PRO nucleic acids
CC can be used for genetic analysis of individuals with genetic disorders.
XX
SQ Sequence 382 AA;

Query Match 99.3%; Score 1986; DB 22; Length 382;
Best Local Similarity 97.9%; Pred. No. 7.2e-183;
Matches 374; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 MRPGTALQAVLLAVLLVGLRAATGRLLS-----GQPVCRGGTQRPCYKVIYFHDTSR 52
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1 MRPGTALQAVLLAVLLVGLRAATGRLLSASDLRLGGQPVCRGGTQRPCYKVIYFHDTSR 60

QY 53 RLNFEEAKEACRRDGGQLVSIIESEDEQKLIKFIENLLPSDGFWIGLRRREEKQSNSTA 112
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
61 RLNFEEAKEACRRDGGQLVSIIESEDEQKLIKFIENLLPSDGFWIGLRRREEKQSNSTA 120

QY 113 CQDLYAWTDGSIQFRNWWYVDEPSCGSEVCVMYHQPAPAGIGGPPYFQWDDRCNMKN 172
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
121 CQDLYAWTDGSIQFRNWWYVDEPSCGSEVCVMYHQPAPAGIGGPPYFQWDDRCNMKN 180

QY 173 NFICKYSDEKPAVPSREAEGETELTTPVLPEETOEDAKTFKESREAAALNLAYILIPS 232
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
181 NFICKYSDEKPAVPSREAEGETELTTPVLPEETOEDAKTFKESREAAALNLAYILIPS 240

QY 233 IPLLLLLLVTTVCVWVICRKRKREQDPDPSTKKQHTTWPSPHQGNSPDLEVYNVIRKQSE 292
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
241 IPLLLLLLVTTVCVWVICRKRKREQDPDPSTKKQHTTWPSPHQGNSPDLEVYNVIRKQSE 300

QY 293 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVEGFTVNDIYE 352
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 301 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVEGFTVNDIYE 360

QY 353 FSPDQMGRSKESGWENEIYGY 374
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
361 FSPDQMGRSKESGWENEIYGY 382

RESULT 7
AAB80235
ID AAB80235 standard; Protein; 382 AA.
XX
AC AAB80235;
XX
DT 24-APR-2001 (first entry)
XX
DE Human PRO234 protein.
XX
KW Human; PRO; dermatological; antipsoriatic; cytostatic; antiinflammatory;
KW antiparkinsonian nootropic; neuroprotective; vulnery; cardiant;
KW antiangiogenic; vasotropic; antiasthmatic; antirheumatic; cancer;
KW antiarthritic; antiinfertility; antidiabetic; antiviral; diabetes;
KW ophthalmological; gene therapy; skin disease; gastrointestinal disorder;
KW ischaemia; inflammation.
XX
OS Homo sapiens.
XX
PN WO200104311-A1.
XX
PD 18-JAN-2001.
XX
PF 22-FEB-2000; 2000WO-US04414.
XX
PR 07-JUL-1999; 99US-0143048.
PR 26-JUL-1999; 99US-0145698.
PR 28-JUL-1999; 99US-0146222.
PR 08-SEP-1999; 99WO-US20594.
PR 13-SEP-1999; 99WO-US20944.
PR 15-SEP-1999; 99WO-US21090.
PR 15-SEP-1999; 99WO-US21547.
PR 05-OCT-1999; 99WO-US23089.
PR 29-NOV-1999; 99WO-US28214.
PR 30-NOV-1999; 99WO-US28313.
PR 16-DEC-1999; 99WO-US30095.
PR 20-DEC-1999; 99WO-US30911.
PR 20-DEC-1999; 99WO-US30999.
PR 05-JAN-2000; 99WO-US00219.
XX
PA (GETH) GENENTECH INC.
XX
PI Ashkenazi AJ, Botstein D, Desnoyers L, Eaton DL, Ferrara N;
PI Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME, Goddard A;
PI Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ, Kljavin IJ;
PI Mather JP, Pan J, Paoni NF, Roy MA, Stewart TA, Tumas D;
PI Williams PM, Wood WI;
XX
DR WPI; 2001-081051/09.
DR N-PSDB; AAF72396.
XX
PT Sixty one nucleic acids encoding PRO polypeptides which are useful in
PT the treatment of skin diseases (e.g. psoriasis), cancers (e.g. lung
PT squamous cell carcinoma) and neurodegenerative diseases (e.g.
PT Alzheimer's disease) -
XX
PS Claim 1; Fig 50; 393pp; English.
XX
CC The present sequence is one of sixty one novel secreted and
CC transmembrane PRO polypeptides. The PRO polypeptides are
CC useful for treating skin diseases (e.g. psoriasis), cancers (e.g. lung
CC squamous cell carcinoma), gastrointestinal disorders (e.g.
CC enterocolitis), neurodegenerative diseases (e.g. Alzheimer's disease,
CC Parkinson's disease), wound repair, cardiovascular disorders (e.g.
CC endometrial bleeding angiogenesis, ischaemias such as coronary
CC ischaemia, atherosclerosis), inflammatory disorders (e.g. asthma,

CC	rheumatoid arthritis, multiple sclerosis), infertility, AIDS and
CC	diabetes and retinal disorders such as retinitis pigmentosum.
CC	The PRO nucleic acids have applications in molecular biology, including
CC	use as hybridization probes, and in chromosome and gene mapping.
XX	
SQ	Sequence 382 AA;
Query Match 99.3%; Score 1986; DB 22; Length 382;	
Best Local Similarity 97.9%; Pred. No. 7.2e-183;	
Matches 374; Conservative 0; Mismatches 0; Indels 8; Gaps 1;	
QY	1 MRPGTALQAVLLAVLLVGLRAATGRLLS-----GQPVCRGGTQRPCVKYIFYFHDTSR 52
Db	
Db	1 MRPGTALQAVLLAVLLVGLRAATGRLLSASDLRGGQPVCRGGTQRPCVKYIFYFHDTSR 60
QY	53 RLNFEEAKEACRRDGGQLVSIIESEDEQKLIIEKFIENLLPSDGFWIGLRRREKQSNSTA 112
Db	
Db	61 RLNFEEAKEACRRDGGQLVSIIESEDEQKLIIEKFIENLLPSDGFWIGLRRREKQSNSTA 120
QY	113 CQDLYAWTDGSIISQFRNWWYVDEPSCGSEVCVVMYHQPSAPAGIGGPFMFQWDDRCNMKN 172
Db	
Db	121 CQDLYAWTDGSIISQFRNWWYVDEPSCGSEVCVVMYHQPSAPAGIGGPFMFQWDDRCNMKN 180
QY	173 NFICKYSDEKPAVPSREAEGETELTTPVLPEETOEDAKKTFKESREALNLAYILIPS 232
Db	
Db	181 NFICKYSDEKPAVPSREAEGETELTTPVLPEETOEDAKKTFKESREALNLAYILIPS 240
QY	233 IPLLLLLLVTTVVVWVICRKRKREQDPSTKKQHTIWPSPHQNSPDLEVYNVIRKQSE 292
Db	
Db	241 IPLLLLLLVTTVVVWVICRKRKREQDPSTKKQHTIWPSPHQNSPDLEVYNVIRKQSE 300
QY	293 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYE 352
Db	
Db	301 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYE 360
QY	353 FSPDQMGRSKESGWVENEIYGY 374
Db	
Db	361 FSPDQMGRSKESGWVENEIYGY 382
RESULT 8	
ABU69645	
ID	ABU69645 standard; Protein; 382 AA.
XX	
AC	ABU69645;
XX	
DT	05-JUN-2003 (first entry)
XX	
DE	Novel human secreted and transmembrane protein PRO234.
XX	
KW	Human; secreted and transmembrane protein; gene therapy; psoriasis;
KW	enterocolitis; gastrointestinal ulceration; skin disease;
KW	keratinocyte differentiation; epithelial cancer; Alzheimer's disease;
KW	squamous cell carcinoma; Parkinson's disease; inflammatory disease;
KW	amyotrophic lateral sclerosis; rheumatoid arthritis; asthma;
KW	multiple sclerosis; organ failure; atherosclerosis; cardiac injury;
KW	infertility; birth defect; premature aging; AIDS; cancer;
KW	diabetic complication; wound repair; tissue re-growth.
OS	Homo sapiens.
XX	
PN	US2003017463-A1.
XX	
PD	23-JAN-2003.
XX	
PF	11-JUL-2001; 2001US-0903640.
XX	
PR	10-SEP-1998; 98WO-US18824.
PR	14-SEP-1998; 98WO-US19177.
PR	16-SEP-1998; 98WO-US19330.
PR	17-SEP-1998; 98WO-US19437.
PR	01-DEC-1998; 98WO-US25108.
PR	08-SEP-1999; 99WO-US20594.

PR	13-SEP-1999;	99WO-US20944.
PR	15-SEP-1999;	99WO-US21090.
PR	15-SEP-1999;	99WO-US21547.
PR	05-OCT-1999;	99WO-US23089.
PR	29-NOV-1999;	99WO-US28214.
PR	30-NOV-1999;	99WO-US28313.
PR	01-DEC-1999;	99WO-US28301.
PR	02-DEC-1999;	99WO-US28564.
PR	02-DEC-1999;	99WO-US28565.
PR	16-DEC-1999;	99WO-US30095.
PR	20-DEC-1999;	99WO-US30911.
PR	20-DEC-1999;	99WO-US30999.
PR	05-JAN-2000;	2000WO-US00219.
PR	11-FEB-2000;	2000WO-US03565.
PR	22-FEB-2000;	2000WO-US04414.
PR	24-FEB-2000;	2000WO-US05004.
PR	02-MAR-2000;	2000WO-US05841.
PR	20-MAR-2000;	2000WO-US07377.
PR	30-MAR-2000;	2000WO-US08439.
PR	22-MAY-2000;	2000WO-US14042.
PR	02-JUN-2000;	2000WO-US15264.
PR	28-JUL-2000;	2000WO-US20710.
PR	24-AUG-2000;	2000WO-US23328.
PR	17-SEP-1997;	97US-059113P.
PR	17-SEP-1997;	97US-059115P.
PR	17-SEP-1997;	97US-059117P.
PR	17-SEP-1997;	97US-059119P.
PR	17-SEP-1997;	97US-059121P.
PR	17-SEP-1997;	97US-059122P.
PR	17-SEP-1997;	97US-059184P.
PR	18-SEP-1997;	97US-059263P.
PR	18-SEP-1997;	97US-059266P.
PR	15-OCT-1997;	97US-062125P.
PR	17-OCT-1997;	97US-062285P.
PR	17-OCT-1997;	97US-062287P.
PR	21-OCT-1997;	97US-063486P.
PR	24-OCT-1997;	97US-062814P.
PR	24-OCT-1997;	97US-062816P.
PR	24-OCT-1997;	97US-063045P.
PR	24-OCT-1997;	97US-063120P.
PR	24-OCT-1997;	97US-063121P.
PR	24-OCT-1997;	97US-063127P.
PR	24-OCT-1997;	97US-063128P.
PR	27-OCT-1997;	97US-063327P.
PR	27-OCT-1997;	97US-063329P.
PR	28-OCT-1997;	97US-063541P.
PR	28-OCT-1997;	97US-063542P.
PR	28-OCT-1997;	97US-063544P.
PR	28-OCT-1997;	97US-063549P.
PR	28-OCT-1997;	97US-063550P.
PR	28-OCT-1997;	97US-063564P.
PR	29-OCT-1997;	97US-063435P.
PR	29-OCT-1997;	97US-063704P.
PR	29-OCT-1997;	97US-063732P.
PR	29-OCT-1997;	97US-063734P.
PR	29-OCT-1997;	97US-063735P.
PR	29-OCT-1997;	97US-063738P.
PR	29-OCT-1997;	97US-064215P.
PR	31-OCT-1997;	97US-063870P.
PR	31-OCT-1997;	97US-064103P.
PR	03-NOV-1997;	97US-064248P.
PR	07-NOV-1997;	97US-064809P.
PR	12-NOV-1997;	97US-065186P.
PR	17-NOV-1997;	97US-065846P.
PR	18-NOV-1997;	97US-065693P.
PR	21-NOV-1997;	97US-066120P.
PR	21-NOV-1997;	97US-066364P.
PR	24-NOV-1997;	97US-066453P.
PR	24-NOV-1997;	97US-066466P.
PR	24-NOV-1997;	97US-066511P.
PR	24-NOV-1997;	97US-066770P.
PR	24-NOV-1997;	97US-066772P.
PR	25-NOV-1997;	97US-066840P.

PR 12-DEC-1997; 97US-069425P.
PR 04-JUN-1998; 98US-088026P.
PR 10-SEP-1998; 98US-099803P.
PR 14-SEP-1998; 98US-100262P.
PR 17-SEP-1998; 98US-100858P.
PR 13-OCT-1998; 98US-104080P.
PR 20-NOV-1998; 98US-109304P.
PR 22-DEC-1998; 98US-113296P.
PR 07-JUL-1999; 99US-143048P.
PR 26-JUL-1999; 99US-145698P.
PR 28-JUL-1999; 99US-146222P.
PR 18-SEP-2000; 2000US-0665350.
XX
PA (GETH) GENENTECH INC.
XX
PI Ashkenazi A, Botstein D, Desnoyers L, Eaton DL, Ferrara N;
PI Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME, Goddard A;
PI Godowski PJ, Grimaldi JC, Gurney AL, Hillan KJ, Kljavin IJ;
PI Mather JP, Pan J, Paoni NF, Roy MA, Stewart TA, Tumas D;
PI Williams PM, Wood WI;
XX
DR WPI; 2003-341586/32.
DR N-PSDB; ACA54920.
XX
PT New PRO polypeptides and nucleic acid molecules, useful in diagnosing
PT or treating inflammatory diseases, organ failure, atherosclerosis,
PT cardiac injury, infertility, cancer, AIDS, Alzheimer's disease or
PT Parkinson's disease -
XX
PS Claim 12; Fig 50; 473pp; English.
XX
CC The invention describes sixty one nucleic acids encoding PRO polypeptides
CC (secreted and transmembrane). The PRO polypeptides and nucleic acids are
CC useful in diagnosing or treating enterocolitis, gastrointestinal
CC ulceration, skin diseases associated with abnormal keratinocyte
CC differentiation, e.g. psoriasis or epithelial cancers such as squamous
CC cell carcinoma, Alzheimer's disease, Parkinson's disease, amyotrophic
CC lateral sclerosis, inflammatory diseases, e.g. rheumatoid arthritis,
CC asthma or multiple sclerosis, organ failure, atherosclerosis, cardiac
CC injury, infertility, birth defects, premature aging, AIDS, cancer,
CC also useful for wound repair and associated therapies concerned with
CC re-growth of tissue. The PRO polypeptides and nucleic acid molecules
CC are also useful in gene therapy, and as molecular weight markers for
CC protein electrophoresis purposes. The anti-PRO antibodies may be used
CC in diagnostic assays for PRO, or for the affinity purification of PRO
CC from recombinant cell culture or natural sources. This is the amino
CC acid sequence of a novel human PRO polypeptide.
XX
SQ Sequence 382 AA;

Query Match 99.3%; Score 1986; DB 24; Length 382;
Best Local Similarity 97.9%; Pred. No. 7.2e-183;
Matches 374; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 MRPGTALQAVLLAVLLVGLRAATGRLLS-----GQVCRGGTQRPCYKVIYFHDTSR 52
Db ||||||||||||||||||||||||||||| |||||||||||||||||||||||
1 MRPGTALQAVLLAVLLVGLRAATGRLLSASDLRLGGQPVCRGGTQRPCYKVIYFHDTSR 60

QY 53 RLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGFWIGLRRREEKQSNSTA 112
Db ||||||||||||||||||||||||||||| |||||||||||||||||||||||
61 RLNFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGFWIGLRRREEKQSNSTA 120

QY 113 CQDLYAWTDGSGISQFRNWWYVDFSCGSEVCVVMYHQPSAPAGIGGPFQWNNDRCNMKN 172
Db ||||||||||||||||||||||||||||| |||||||||||||||||||||||
121 CQDLYAWTDGSGISQFRNWWYVDFSCGSEVCVVMYHQPSAPAGIGGPFQWNNDRCNMKN 180

QY 173 NFICKYSDEKPAVPSRAEAGEETELTTPVLPEETOEDAKKTFKESREALNLAYILIPS 232
Db ||||||||||||||||||||||||||||| |||||||||||||||||||||||
181 NFICKYSDEKPAVPSRAEAGEETELTTPVLPEETOEDAKKTFKESREALNLAYILIPS 240

QY 233 IPLLLLLVVTTCVWVICRKRKREQDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSE 292
|||

Db 241 IPLLLLLVVTTCVWVICRKRKREQDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSE 300
QY 293 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYE 352
||||||||||||||||||||||||||| |||||||||||||||||||||||
Db 301 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYE 360
QY 353 FSPDQMGSRKESGWVENEIYGY 374
|||||||||||||||||||||||
Db 361 FSPDQMGSRKESGWVENEIYGY 382

RESULT 9
ABU71121
ID ABU71121 standard; Protein; 382 AA.
XX
AC ABU71121;
XX
DT 10-JUN-2003 (first entry)
XX
DE Human PRO234 protein.
XX
KW Human; PRO; secreted; transmembrane; cytostatic; TNF-alpha; blood;
KW tumour necrosis factor alpha release; chondrocyte cell; proliferation;
KW differentiation; tumour; gene therapy.
XX
OS Homo sapiens.
XX
PN US2003036143-A1.
XX
PD 20-FEB-2003.
XX
PF 02-JUL-2002; 2002US-0187600.
XX
PR 16-SEP-1998; 98WO-US19330.
PR 07-OCT-1998; 98WO-US21141.
PR 01-DEC-1998; 98WO-US25108.
PR 08-MAR-1999; 99WO-US05028.
PR 14-MAY-1999; 99WO-US10733.
PR 02-JUN-1999; 99WO-US12252.
PR 01-SEP-1999; 99WO-US20111.
PR 15-SEP-1999; 99WO-US21090.
PR 01-DEC-1999; 99WO-US28301.
PR 02-DEC-1999; 99WO-US28551.
PR 30-DEC-1999; 99WO-US31274.
PR 05-JAN-2000; 2000WO-US00219.
PR 18-FEB-2000; 2000WO-US04341.
PR 18-FEB-2000; 2000WO-US04342.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US05004.
PR 01-MAR-2000; 2000WO-US05601.
PR 02-MAR-2000; 2000WO-US05841.
PR 15-MAR-2000; 2000WO-US06884.
PR 30-MAR-2000; 2000WO-US08439.
PR 17-MAY-2000; 2000WO-US13705.
PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.
PR 02-JUN-2000; 2000WO-US15264.
PR 28-JUL-2000; 2000WO-US20710.
PR 24-AUG-2000; 2000WO-US23328.
PR 08-NOV-2000; 2000WO-US30952.
PR 01-DEC-2000; 2000WO-US32678.
PR 20-DEC-2000; 2000WO-US34956.
PR 28-FEB-2001; 2001WO-US06520.
PR 01-JUN-2001; 2001WO-US17800.
PR 20-JUN-2001; 2001WO-US19692.
PR 29-JUN-2001; 2001WO-US21066.
PR 09-JUL-2001; 2001WO-US21735.
PR 29-AUG-2001; 2001WO-US27099.
PR 18-SEP-1997; 97US-059263P.
PR 18-SEP-1997; 97US-059266P.
PR 17-OCT-1997; 97US-062250P.
PR 21-OCT-1997; 97US-063486P.
PR 24-OCT-1997; 97US-063120P.

PR 24-OCT-1997; 97US-063121P.
PR 28-OCT-1997; 97US-063540P.
PR 28-OCT-1997; 97US-063541P.
PR 28-OCT-1997; 97US-063544P.
PR 28-OCT-1997; 97US-063564P.
PR 29-OCT-1997; 97US-063734P.
PR 31-OCT-1997; 97US-063870P.
PR 31-OCT-1997; 97US-064103P.
PR 13-NOV-1997; 97US-065311P.
PR 21-NOV-1997; 97US-066120P.
PR 24-NOV-1997; 97US-066466P.
PR 24-NOV-1997; 97US-066772P.
PR 11-DEC-1997; 97US-069335P.
PR 12-DEC-1997; 97US-069425P.
PR 17-DEC-1997; 97US-069870P.
PR 18-DEC-1997; 97US-068017P.
PR 10-MAR-1998; 98US-077450P.
PR 11-MAR-1998; 98US-077632P.
PR 11-MAR-1998; 98US-077649P.
PR 20-MAR-1998; 98US-078886P.
PR 20-MAR-1998; 98US-078939P.
PR 27-MAR-1998; 98US-079664P.
PR 27-MAR-1998; 98US-079786P.
PR 31-MAR-1998; 98US-080107P.
PR 31-MAR-1998; 98US-080194P.
PR 01-APR-1998; 98US-080327P.
PR 01-APR-1998; 98US-080333P.
PR 08-APR-1998; 98US-081049P.
PR 08-APR-1998; 98US-081070P.
PR 09-APR-1998; 98US-081195P.
PR 15-APR-1998; 98US-081838P.
PR 21-APR-1998; 98US-082568P.
PR 21-APR-1998; 98US-082569P.
PR 22-APR-1998; 98US-082704P.
PR 22-APR-1998; 98US-082797P.
PR 28-APR-1998; 98US-083322P.
PR 29-APR-1998; 98US-083495P.
PR 29-APR-1998; 98US-083496P.
PR 29-APR-1998; 98US-083499P.
PR 29-APR-1998; 98US-083559P.
PR 05-MAY-1998; 98US-084366P.
PR 06-MAY-1998; 98US-084414P.
PR 07-MAY-1998; 98US-084639P.
PR 07-MAY-1998; 98US-084640P.
PR 07-MAY-1998; 98US-084643P.
PR 15-MAY-1998; 98US-085579P.
PR 15-MAY-1998; 98US-085580P.
PR 15-MAY-1998; 98US-085582P.
PR 15-MAY-1998; 98US-085700P.
PR 18-MAY-1998; 98US-086023P.
PR 22-MAY-1998; 98US-086392P.
PR 22-MAY-1998; 98US-086486P.
PR 28-MAY-1998; 98US-087098P.
PR 28-MAY-1998; 98US-087208P.
PR 02-JUN-1998; 98US-087609P.
PR 02-JUN-1998; 98US-087759P.
PR 03-JUN-1998; 98US-088025P.
PR 04-JUN-1998; 98US-088028P.
PR 04-JUN-1998; 98US-088029P.
PR 04-JUN-1998; 98US-088033P.
PR 04-JUN-1998; 98US-088326P.
PR 05-JUN-1998; 98US-088167P.
PR 05-JUN-1998; 98US-088202P.
PR 05-JUN-1998; 98US-088212P.
PR 05-JUN-1998; 98US-088217P.
PR 09-JUN-1998; 98US-088655P.
PR 10-JUN-1998; 98US-088722P.
PR 10-JUN-1998; 98US-088738P.
PR 10-JUN-1998; 98US-088740P.
PR 10-JUN-1998; 98US-088811P.
PR 10-JUN-1998; 98US-088824P.
PR 10-JUN-1998; 98US-088825P.

PR 10-JUN-1998; 98US-088826P.
PR 11-JUN-1998; 98US-088861P.
PR 11-JUN-1998; 98US-088863P.
PR 11-JUN-1998; 98US-088876P.
PR 12-JUN-1998; 98US-089090P.
PR 12-JUN-1998; 98US-089105P.
PR 16-JUN-1998; 98US-089512P.
PR 16-JUN-1998; 98US-089514P.
PR 17-JUN-1998; 98US-089538P.
PR 17-JUN-1998; 98US-089598P.
PR 17-JUN-1998; 98US-089653P.
PR 18-JUN-1998; 98US-089908P.
PR 19-JUN-1998; 98US-089952P.
PR 22-JUN-1998; 98US-090246P.
PR 22-JUN-1998; 98US-090252P.
PR 22-JUN-1998; 98US-090254P.
PR 24-JUN-1998; 98US-090429P.
PR 24-JUN-1998; 98US-090435P.
PR 24-JUN-1998; 98US-090444P.
PR 24-JUN-1998; 98US-090461P.
PR 24-JUN-1998; 98US-090535P.
PR 24-JUN-1998; 98US-090540P.
PR 25-JUN-1998; 98US-090676P.
PR 25-JUN-1998; 98US-090678P.
PR 25-JUN-1998; 98US-090688P.
PR 25-JUN-1998; 98US-090690P.
PR 25-JUN-1998; 98US-090694P.
PR 25-JUN-1998; 98US-090695P.
PR 25-JUN-1998; 98US-090696P.
PR 26-JUN-1998; 98US-090862P.
PR 26-JUN-1998; 98US-090863P.
PR 26-JUN-1998; 98US-091010P.
PR 01-JUL-1998; 98US-091359P.
PR 01-JUL-1998; 98US-091544P.
PR 02-JUL-1998; 98US-091478P.
PR 02-JUL-1998; 98US-091486P.
PR 02-JUL-1998; 98US-091626P.
PR 02-JUL-1998; 98US-091628P.
PR 02-JUL-1998; 98US-091632P.
PR 24-JUL-1998; 98US-094006P.
PR 04-AUG-1998; 98US-095282P.
PR 10-AUG-1998; 98US-095998P.
PR 10-AUG-1998; 98US-096012P.
PR 17-AUG-1998; 98US-096757P.
PR 17-AUG-1998; 98US-096766P.
PR 17-AUG-1998; 98US-096867P.
PR 17-AUG-1998; 98US-096891P.
PR 17-AUG-1998; 98US-096897P.
PR 18-AUG-1998; 98US-096949P.
PR 18-AUG-1998; 98US-096959P.
PR 18-AUG-1998; 98US-097022P.
PR 26-AUG-1998; 98US-097952P.
PR 26-AUG-1998; 98US-097954P.
PR 26-AUG-1998; 98US-097955P.
PR 26-AUG-1998; 98US-097971P.
PR 26-AUG-1998; 98US-097974P.
PR 26-AUG-1998; 98US-098014P.
PR 01-SEP-1998; 98US-098716P.
PR 01-SEP-1998; 98US-098723P.
PR 02-SEP-1998; 98US-098803P.
PR 02-SEP-1998; 98US-098821P.
PR 02-SEP-1998; 98US-098843P.
PR 09-SEP-1998; 98US-099602P.
PR 10-SEP-1998; 98US-099741P.
PR 10-SEP-1998; 98US-099754P.
PR 10-SEP-1998; 98US-099763P.
PR 10-SEP-1998; 98US-099812P.

Query Match 99.3%; Score 1986; DB 24; Length 382;
Best Local Similarity 97.9%; Pred. No. 7.2e-183;
Matches 374; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 MRPGTALQAVLLAVLLVGLRAATGRLLS-----GQPVCRGGTQPCYKVIYFHDTSR 52


```
Db      ||||| 1 MRPGTALQAVLLAVLLVGLRAATGRLLSASDLDLRGGQPVCRGGTQPCYKVIYFHDTSR 60
Qy      53 RLNFEEAKEACRRDGGQLVSIIESEDEQKLEKFIENLLPSDGDGFWIGLRRRREEKQSNSTA 112
Db      61 RLNFEEAKEACRRDGGQLVSIIESEDEQKLEKFIENLLPSDGDGFWIGLRRRREEKQSNSTA 120
Qy      113 CQDLYAWTDGSIQFRNWWYVDEPSCGSEVCVVMYHOPSAFAGIGGPPYMFQWDDRCNMKN 172
Db      121 CQDLYAWTDGSIQFRNWWYVDEPSCGSEVCVVMYHOPSAFAGIGGPPYMFQWDDRCNMKN 180
Qy      173 NFICKYSDEKPAVPSREAEGETELTPVLPEETQEEADAKTKFESREAAALNLAYILIPS 232
Db      181 NFICKYSDEKPAVPSREAEGETELTPVLPEETQEEADAKTKFESREAAALNLAYILIPS 240
Qy      233 IPLLALLVVTTCVWVICRKRKREQDPSTKKQHTIWPSPHQNSPDLEVYNVIRKQSE 292
Db      241 IPLLALLVVTTCVWVICRKRKREQDPSTKKQHTIWPSPHQNSPDLEVYNVIRKQSE 300
Qy      293 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFTNDIYE 352
Db      301 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFTNDIYE 360
Qy      353 FSPDQMGSRKESGWVENEIYG 374
Db      361 FSPDQMGSRKESGWVENEIYG 382
```

```
RESULT 10
ABU71468
ID      ABU71468 standard; Protein; 382 AA.
XX      ABU71468;
XX      10-JUN-2003 (first entry)
DE      Human PRO polypeptide #24.
XX      Human; secreted and transmembrane protein; PRO polypeptide; cancer;
KW      Alzheimer's disease; ischaemia; cytostatic; nootropic; vasotropic;
KW      neuroprotective.
XX      Homo sapiens.
OS      US2002192659-A1.
PN      19-DEC-2002.
XX      10-JUL-2001; 2001US-0902853.
PF      10-SEP-1998; 98WO-US18824.
XX      14-SEP-1998; 98WO-US19177.
PR      16-SEP-1998; 98WO-US19330.
PR      17-SEP-1998; 98WO-US19437.
PR      01-DEC-1998; 98WO-US25108.
PR      08-SEP-1999; 99WO-US20594.
PR      13-SEP-1999; 99WO-US20944.
PR      15-SEP-1999; 99WO-US21090.
PR      15-SEP-1999; 99WO-US21547.
PR      05-OCT-1999; 99WO-US23089.
PR      01-DEC-1999; 99WO-US28301.
PR      02-DEC-1999; 99WO-US28564.
PR      02-DEC-1999; 99WO-US28565.
PR      16-DEC-1999; 99WO-US30095.
PR      20-DEC-1999; 99WO-US30911.
PR      20-DEC-1999; 99WO-US30999.
PR      05-JAN-2000; 2000WO-US00219.
PR      11-FEB-2000; 2000WO-US03565.
PR      22-FEB-2000; 2000WO-US04414.
PR      28-JUL-2000; 2000WO-US20710.
PR      24-AUG-2000; 2000WO-US23328.
PR      17-SEP-1997; 97US-059113P.
PR      17-SEP-1997; 97US-059115P.
```

```
PR      17-SEP-1997; 97US-059117P.
PR      18-SEP-1997; 97US-059266P.
PR      15-OCT-1997; 97US-062125P.
PR      17-OCT-1997; 97US-062285P.
PR      17-OCT-1997; 97US-062287P.
PR      21-OCT-1997; 97US-063486P.
PR      24-OCT-1997; 97US-062814P.
PR      24-OCT-1997; 97US-062816P.
XX      (GETH ) GENENTECH INC.
XX      Ashkenazi A, Botstein D, Desnoyers L, Eaton DL, Ferrara N;
PI      Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME, Goddard A;
PI      Godowski PJ, Grimaldi JC, Gurney AL, Hillan KJ, Kijavir IJ;
PI      Mather JP, Pan J, Paoni NF, Roy MA, Stewart TA, Tumas D;
PI      Williams PM, Wood WI;
XX      WPI; 2003-361832/34.
DR      N-PSDB; ACA58405.
XX      New isolated nucleic acid encoding a PRO polypeptide, e.g. PRO245 or
PT      PRO1868, useful in molecular biology, chromosome and gene mapping, in
PT      generating antisense RNA and DNA, and in gene therapy -
XX      Claim 12; Fig 50; 474pp; English.
XX      The present invention relates to the isolation of novel human secreted
CC      and transmembrane proteins (PRO polypeptides), and the polynucleotide
CC      sequences encoding them. The polynucleotide sequences are useful in
CC      molecular biology, as hybridisation probes, in chromosome and gene
CC      mapping, in generating antisense RNA and DNA, and in gene therapy. The
CC      polynucleotide sequences may also be used in preparing PRO polypeptides
CC      by recombinant techniques, and in generating either transgenic animals
CC      or knock-out animals which, in turn, are useful in the development and
CC      screening of therapeutically useful reagents. The PRO polypeptides or
CC      their antibodies are useful in preparing a medicament for treating a
CC      condition responsive to the polypeptide or antibody, such as cancer,
CC      Alzheimer's disease or ischaemia, and in various diagnostic assays.
CC      ABU71445-ABU71505 represent human PRO polypeptides of the invention.
XX      SQ      Sequence 382 AA;
XX      Query Match 99.3%; Score 1986; DB 24; Length 382;
XX      Best Local Similarity 97.9%; Pred. No. 7.2e-183;
XX      Matches 374; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
Qy      1 MRPGTALQAVLLAVLLVGLRAATGRLLS-----GQPVCRGGTQPCYKVIYFHDTSR 52
Db      1 MRPGTALQAVLLAVLLVGLRAATGRLLSASDLDLRGGQPVCRGGTQPCYKVIYFHDTSR 60
Qy      53 RLNFEEAKEACRRDGGQLVSIIESEDEQKLEKFIENLLPSDGDGFWIGLRRRREEKQSNSTA 112
Db      61 RLNFEEAKEACRRDGGQLVSIIESEDEQKLEKFIENLLPSDGDGFWIGLRRRREEKQSNSTA 120
Qy      113 CQDLYAWTDGSIQFRNWWYVDEPSCGSEVCVVMYHOPSAFAGIGGPPYMFQWDDRCNMKN 172
Db      121 CQDLYAWTDGSIQFRNWWYVDEPSCGSEVCVVMYHOPSAFAGIGGPPYMFQWDDRCNMKN 180
Qy      173 NFICKYSDEKPAVPSREAEGETELTPVLPEETQEEADAKTKFESREAAALNLAYILIPS 232
Db      181 NFICKYSDEKPAVPSREAEGETELTPVLPEETQEEADAKTKFESREAAALNLAYILIPS 240
Qy      233 IPLLALLVVTTCVWVICRKRKREQDPSTKKQHTIWPSPHQNSPDLEVYNVIRKQSE 292
Db      241 IPLLALLVVTTCVWVICRKRKREQDPSTKKQHTIWPSPHQNSPDLEVYNVIRKQSE 300
Qy      293 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFTNDIYE 352
Db      301 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFTNDIYE 360
Qy      353 FSPDQMGSRKESGWVENEIYG 374
Db      361 FSPDQMGSRKESGWVENEIYG 382
```

RESULT 11
ABU71914
ID ABU71914 standard; Protein; 382 AA.
XX
AC ABU71914;
XX
DT 12-JUN-2003 (first entry)
XX
DE Human secreted/transmembrane protein PRO234.
XX
KW Human; secreted protein; transmembrane protein; PRO;
KW gene therapy; chromosome identification; chromosome marker.
XX
OS Homo sapiens.
XX
PN US2003003530-A1.
XX
PD (2-JAN-2003.
XX
PF 11-JUL-2001; 2001US-0904011.
XX
PR 10-SEP-1998; 98WO-US18824.
PR 14-SEP-1998; 98WO-US19177.
PR 16-SEP-1998; 98WO-US19330.
PR 17-SEP-1998; 98WO-US19437.
PR 01-DEC-1998; 98WO-US25108.
PR 08-SEP-1999; 99WO-US20594.
PR 13-SEP-1999; 99WO-US20944.
PR 15-SEP-1999; 99WO-US21090.
PR 05-OCT-1999; 99WO-US21547.
PR 29-NOV-1999; 99WO-US23089.
PR 30-NOV-1999; 99WO-US28214.
PR 01-DEC-1999; 99WO-US28313.
PR 02-DEC-1999; 99WO-US28564.
PR 02-DEC-1999; 99WO-US28565.
PR 16-DEC-1999; 99WO-US30095.
PR 20-DEC-1999; 99WO-US30911.
PR 20-DEC-1999; 99WO-US30999.
PR 05-JAN-2000; 2000WO-US00219.
PR 11-FEB-2000; 2000WO-US03565.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US05004.
PR 02-MAR-2000; 2000WO-US05841.
PR 20-MAR-2000; 2000WO-US07377.
PR 30-MAR-2000; 2000WO-US08439.
PR 22-MAY-2000; 2000WO-US14042.
PR 02-JUN-2000; 2000WO-US15264.
PR 28-JUL-2000; 2000WO-US20710.
PR 24-AUG-2000; 2000WO-US23328.
PR 17-SEP-1997; 97US-059113P.
PR 17-SEP-1997; 97US-059115P.
PR 17-SEP-1997; 97US-059117P.
PR 17-SEP-1997; 97US-059119P.
PR 17-SEP-1997; 97US-059121P.
PR 17-SEP-1997; 97US-059122P.
PR 17-SEP-1997; 97US-059184P.
PR 18-SEP-1997; 97US-059263P.
PR 18-SEP-1997; 97US-059266P.
PR 15-OCT-1997; 97US-062125P.
PR 17-OCT-1997; 97US-062285P.
PR 17-OCT-1997; 97US-062287P.
PR 21-OCT-1997; 97US-063486P.
PR 24-OCT-1997; 97US-062814P.
PR 24-OCT-1997; 97US-062816P.
PR 24-OCT-1997; 97US-063045P.
PR 24-OCT-1997; 97US-063120P.
PR 24-OCT-1997; 97US-063121P.
PR 24-OCT-1997; 97US-063127P.
PR 24-OCT-1997; 97US-063128P.
PR 27-OCT-1997; 97US-063327P.

PR 27-OCT-1997; 97US-063329P.
PR 28-OCT-1997; 97US-063541P.
PR 28-OCT-1997; 97US-063542P.
PR 28-OCT-1997; 97US-063544P.
PR 28-OCT-1997; 97US-063549P.
PR 28-OCT-1997; 97US-063550P.
PR 28-OCT-1997; 97US-063564P.
PR 29-OCT-1997; 97US-063435P.
PR 29-OCT-1997; 97US-063704P.
PR 29-OCT-1997; 97US-063732P.
PR 29-OCT-1997; 97US-063734P.
PR 29-OCT-1997; 97US-063735P.
PR 29-OCT-1997; 97US-063738P.
PR 29-OCT-1997; 97US-064215P.
PR 31-OCT-1997; 97US-063870P.
PR 31-OCT-1997; 97US-064103P.
PR 03-NOV-1997; 97US-064248P.
PR 07-NOV-1997; 97US-064809P.
PR 12-NOV-1997; 97US-065186P.
PR 17-NOV-1997; 97US-065846P.
PR 18-NOV-1997; 97US-065693P.
PR 21-NOV-1997; 97US-066120P.
PR 21-NOV-1997; 97US-066364P.
PR 24-NOV-1997; 97US-066453P.
PR 24-NOV-1997; 97US-066466P.
PR 24-NOV-1997; 97US-066511P.
PR 24-NOV-1997; 97US-066770P.
PR 24-NOV-1997; 97US-066772P.
PR 18-SEP-2000; 2000US-0665350.
XX
PA (GETH) GENENTECH INC.
XX
PI Ashkenazi A, Botstein D, Desnoyers L, Eaton DL, Ferrara N;
PI Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME, Goddard A;
PI Godowski PJ, Grimaldi JC, Gurney AL, Hillan KJ, Kljavin IJ;
PI Mather JP, Pan J, Paoni NF, Roy MA, Stewart TA, Tumas D;
PI Williams PM, Wood WI;
XX
DR WPI; 2003-329602/31.
DR N-PSDB; ACA60112.
XX
PT New transmembrane polypeptides and nucleic acids encoding the
PT polypeptides, useful in gene therapy, in chromosome identification, as
PT chromosome markers, in generating probes and in tissue typing -
XX
PS Claim 12; Fig 50; 484pp; English.
XX
CC The invention relates to an isolated nucleic acid with at least 80%
CC nucleic acid sequence identity to a nucleotide sequence encoding one of
CC 61 secreted/transmembrane polypeptides, or PRO polypeptides or encoding a
CC PRO protein extracellular domain. Also included are a vector comprising
CC the PRO nucleic acid, a host cell comprising the vector, producing a PRO
CC polypeptide (by culturing the host cell for the expression of the PRO
CC polypeptide, and recovering the PRO polypeptide from the cell culture),
CC an isolated PRO polypeptide (having at least 80% sequence identity
CC to: (a) an amino acid sequence selected from the 61 PRO proteins;
CC (b) an amino acid sequence encoded by a nucleic acid molecule deposited
CC with an ATCC number (detailed in the specification); or (c) an
CC extracellular domain of a PRO polypeptide or to a PRO polypeptide lacking
CC its associated signal peptide), a chimeric molecule comprising a PRO
CC polypeptide of fused to a heterologous amino acid sequence, an anti-PRO
CC antibody, detecting a PRO245 or PRO1868 in a sample suspected of
CC containing the polypeptide, linking a bioactive molecule to a cell
CC expressing a PRO245 or PRO1868 and modulating at least one biological
CC activity of a cell expressing a PRO245 or PRO1868. Nucleic acids which
CC encode PRO can be used to generate either transgenic animals or knock-out
CC animals which may be used in the development and screening of
CC therapeutically useful reagents. The nucleic acids may also be used in
CC gene therapy, in chromosome identification, as chromosome markers, or in
CC generating probes. The PRO polypeptides are useful as molecular markers
CC for protein electrophoresis, and the isolated nucleic acids may be used
CC for recombinantly expressing those markers. The PRO polypeptides and
CC nucleic acids may also be used in tissue typing. Anti-PRO antibodies

CC are useful in diagnostic assays for PRO, and in affinity purification
CC of PRO from recombinant cell culture or natural sources. The
CC present sequence represents a PRO protein.

XX
SQ Sequence 382 AA;

Query Match 99.3%; Score 1986; DB 24; Length 382;
Best Local Similarity 97.9%; Pred. No. 7.2e-183;
Matches 374; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 MRPGTALQAVLLAVLLVGLRAATGRLLS-----GQVCRGGTQRPCYKVIYFHDTSR 52
Db 1 MRPGTALQAVLLAVLLVGLRAATGRLLSASDLDRGGQPVCRGGTQRPCYKVIYFHDTSR 60

QY 53 RLNFEEAKEACRRDGGQLVSIIESEDEQKLIIEKFIENLLPSDGFWIGLRRREKQSNSTA 112
Db 61 RLNFEEAKEACRRDGGQLVSIIESEDEQKLIIEKFIENLLPSDGFWIGLRRREKQSNSTA 120

QY 113 QDLYAWTDGSIQFRNMYVDEPSCGSEVCVMYHQPSAPAGIGGYPMFQWDDRCNMKN 172
Db 121 QDLYAWTDGSIQFRNMYVDEPSCGSEVCVMYHQPSAPAGIGGYPMFQWDDRCNMKN 180

QY 173 NFICKYSDEKPAVPSREAEGETELTTPVLPEETQEDAKKTFKESREAAINLAYILIPS 232
Db 181 NFICKYSDEKPAVPSREAEGETELTTPVLPEETQEDAKKTFKESREAAINLAYILIPS 240

QY 233 IPLLALLLVTTVVCWWICRKRKREQDPDPSTKKQHTIWPSPHQGNSPDLEYNVIRKQSE 292
Db 241 IPLLALLLVTTVVCWWICRKRKREQDPDPSTKKQHTIWPSPHQGNSPDLEYNVIRKQSE 300

QY 293 ADLAETRPDLKNISFRVCSGEATPDDMSCDYNMNAVNPSESGFVTLVSVEGFTNDIYE 352
Db 301 ADLAETRPDLKNISFRVCSGEATPDDMSCDYNMNAVNPSESGFVTLVSVEGFTNDIYE 360

QY 353 FSPDQMGRSKESGWVENEIYGY 374
Db 361 FSPDQMGRSKESGWVENEIYGY 382

RESULT 12
ABU65578
ID ABU65578 standard; Protein; 382 AA.
XX
AC ABU65578;
XX
DT 19-MAY-2003 (first entry)
XX
DE Human secreted/transmembrane protein, SEQ ID 20.
XX
KW Human; PRO; secreted protein; transmembrane protein;
KW cytosolic; antiarthritic; osteopathic; adrenal tumour; lung tumour;
KW colon tumour; breast tumour; prostate tumour; rectal tumour;
KW cervical tumour; liver tumour; TNF-alpha release; arthritis;
KW tumour necrosis factor alpha; chondrocyte cell; bone disorder;
KW cartilage disorder; sports injury.
XX
OS Homo sapiens.
XX
PN US2003036156-A1.
XX
PD 20-FEB-2003.
XX
PF 02-JUL-2002; 2002US-0188767.
XX
PR 16-SEP-1998; 98WO-US19330.
PR 07-OCT-1998; 98WO-US21141.
PR 01-DEC-1998; 98WO-US25108.
PR 08-MAR-1999; 99WO-US05028.
PR 14-MAY-1999; 99WO-US10733.
PR 02-JUN-1999; 99WO-US12252.
PR 01-SEP-1999; 99WO-US20111.
PR 15-SEP-1999; 99WO-US21090.
PR 01-DEC-1999; 99WO-US28301.

PR 02-DEC-1999; 99WO-US28551.
PR 30-DEC-1999; 99WO-US31274.
PR 05-JAN-2000; 2000WO-US00219.
PR 18-FEB-2000; 2000WO-US04341.
PR 18-FEB-2000; 2000WO-US04342.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US05004.
PR 01-MAR-2000; 2000WO-US05601.
PR 02-MAR-2000; 2000WO-US05841.
PR 15-MAR-2000; 2000WO-US06884.
PR 30-MAR-2000; 2000WO-US08439.
PR 17-MAY-2000; 2000WO-US13705.
PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.
PR 02-JUN-2000; 2000WO-US15264.
PR 28-JUL-2000; 2000WO-US20710.
PR 24-AUG-2000; 2000WO-US23328.
PR 08-NOV-2000; 2000WO-US30952.
PR 01-DEC-2000; 2000WO-US32678.
PR 20-DEC-2000; 2000WO-US34956.
PR 28-FEB-2001; 2001WO-US06520.
PR 01-JUN-2001; 2001WO-US17800.
PR 20-JUN-2001; 2001WO-US19692.
PR 29-JUN-2001; 2001WO-US21066.
PR 09-JUL-2001; 2001WO-US21735.
PR 29-AUG-2001; 2001WO-US27099.
PR 18-SEP-1997; 97US-059263P.
PR 18-SEP-1997; 97US-059266P.
PR 17-OCT-1997; 97US-062250P.
PR 21-OCT-1997; 97US-063486P.
PR 24-OCT-1997; 97US-063120P.
PR 24-OCT-1997; 97US-063121P.
PR 28-OCT-1997; 97US-063540P.
PR 28-OCT-1997; 97US-063541P.
PR 28-OCT-1997; 97US-063544P.
PR 28-OCT-1997; 97US-063564P.
PR 29-OCT-1997; 97US-063734P.
PR 31-OCT-1997; 97US-063870P.
PR 31-OCT-1997; 97US-064103P.
PR 13-NOV-1997; 97US-065311P.
PR 21-NOV-1997; 97US-066120P.
PR 24-NOV-1997; 97US-066466P.
PR 24-NOV-1997; 97US-066772P.
PR 11-DEC-1997; 97US-069335P.
PR 12-DEC-1997; 97US-069425P.
PR 17-DEC-1997; 97US-069870P.
PR 18-DEC-1997; 97US-068017P.
PR 10-MAR-1998; 98US-077450P.
PR 11-MAR-1998; 98US-077632P.
PR 11-MAR-1998; 98US-077649P.
PR 20-MAR-1998; 98US-078886P.
PR 20-MAR-1998; 98US-078939P.
PR 27-MAR-1998; 98US-079664P.
PR 27-MAR-1998; 98US-079786P.
PR 31-MAR-1998; 98US-080107P.
PR 31-MAR-1998; 98US-080194P.
PR 01-APR-1998; 98US-080327P.
PR 01-APR-1998; 98US-080333P.
PR 08-APR-1998; 98US-081049P.
PR 08-APR-1998; 98US-081070P.
PR 09-APR-1998; 98US-081195P.
PR 15-APR-1998; 98US-081838P.
PR 21-APR-1998; 98US-082568P.
PR 21-APR-1998; 98US-082569P.
PR 22-APR-1998; 98US-082704P.
PR 22-APR-1998; 98US-082797P.
PR 28-APR-1998; 98US-083322P.
PR 29-APR-1998; 98US-083495P.
PR 29-APR-1998; 98US-083496P.
PR 29-APR-1998; 98US-083499P.
PR 29-APR-1998; 98US-083559P.
PR 05-MAY-1998; 98US-084366P.
PR 06-MAY-1998; 98US-084414P.

PN US2003036157-A1.
XX 20-FEB-2003.
PD
XX
PF 02-JUL-2002; 2002US-0188769.
XX
PR 16-SEP-1998; 98WO-US19330.
PR 07-OCT-1998; 98WO-US21141.
PR 01-DEC-1998; 98WO-US25108.
PR 08-MAR-1999; 99WO-US05028.
PR 14-MAY-1999; 99WO-US10733.
PR 02-JUN-1999; 99WO-US12252.
PR 01-SEP-1999; 99WO-US20111.
PR 15-SEP-1999; 99WO-US21090.
PR 01-DEC-1999; 99WO-US28301.
PR 02-DEC-1999; 99WO-US28551.
PR 30-DEC-1999; 99WO-US31274.
PR 05-JAN-2000; 2000WO-US00219.
PR 18-FEB-2000; 2000WO-US04341.
PR 18-FEB-2000; 2000WO-US04342.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US05004.
PR 01-MAR-2000; 2000WO-US05601.
PR 02-MAR-2000; 2000WO-US05841.
PR 15-MAR-2000; 2000WO-US06884.
PR 30-MAR-2000; 2000WO-US08439.
PR 17-MAY-2000; 2000WO-US13705.
PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.
PR 02-JUN-2000; 2000WO-US15264.
PR 28-JUL-2000; 2000WO-US20710.
PR 24-AUG-2000; 2000WO-US23328.
PR 08-NOV-2000; 2000WO-US30952.
PR 01-DEC-2000; 2000WO-US32678.
PR 20-DEC-2000; 2000WO-US34956.
PR 28-FEB-2001; 2001WO-US06520.
PR 01-JUN-2001; 2001WO-US17800.
PR 20-JUN-2001; 2001WO-US19692.
PR 29-JUN-2001; 2001WO-US21066.
PR 09-JUL-2001; 2001WO-US21735.
PR 29-AUG-2001; 2001WO-US27099.
PR 18-SEP-1997; 97US-059263P.
PR 18-SEP-1997; 97US-059266P.
PR 17-OCT-1997; 97US-062250P.
PR 21-OCT-1997; 97US-063486P.
PR 24-OCT-1997; 97US-063120P.
PR 24-OCT-1997; 97US-063121P.
PR 28-OCT-1997; 97US-063540P.
PR 28-OCT-1997; 97US-063541P.
PR 28-OCT-1997; 97US-063544P.
PR 28-OCT-1997; 97US-063564P.
PR 29-OCT-1997; 97US-063734P.
PR 31-OCT-1997; 97US-063870P.
PR 31-OCT-1997; 97US-064103P.
PR 13-NOV-1997; 97US-065311P.
PR 21-NOV-1997; 97US-066120P.
PR 24-NOV-1997; 97US-066466P.
PR 24-NOV-1997; 97US-066772P.
PR 11-DEC-1997; 97US-069335P.
PR 12-DEC-1997; 97US-069425P.
PR 17-DEC-1997; 97US-069870P.
PR 18-DEC-1997; 97US-068017P.
PR 10-MAR-1998; 98US-077450P.
PR 11-MAR-1998; 98US-077632P.
PR 11-MAR-1998; 98US-077649P.
PR 20-MAR-1998; 98US-07886P.
PR 20-MAR-1998; 98US-078939P.
PR 27-MAR-1998; 98US-079664P.
PR 27-MAR-1998; 98US-079786P.
PR 31-MAR-1998; 98US-080107P.
PR 31-MAR-1998; 98US-080194P.
PR 01-APR-1998; 98US-080327P.
PR 01-APR-1998; 98US-080333P.

PR 08-APR-1998; 98US-081049P.
PR 08-APR-1998; 98US-081070P.
PR 09-APR-1998; 98US-081195P.
PR 15-APR-1998; 98US-081838P.
PR 21-APR-1998; 98US-082568P.
PR 21-APR-1998; 98US-082569P.
PR 22-APR-1998; 98US-082704P.
PR 22-APR-1998; 98US-082797P.
PR 28-APR-1998; 98US-083322P.
PR 29-APR-1998; 98US-083495P.
PR 29-APR-1998; 98US-083496P.
PR 29-APR-1998; 98US-083499P.
PR 29-APR-1998; 98US-083559P.
PR 05-MAY-1998; 98US-084366P.
PR 06-MAY-1998; 98US-084414P.
PR 07-MAY-1998; 98US-084639P.
PR 07-MAY-1998; 98US-084640P.
PR 07-MAY-1998; 98US-084643P.
PR 15-MAY-1998; 98US-085579P.
PR 15-MAY-1998; 98US-085580P.
PR 15-MAY-1998; 98US-085582P.
PR 15-MAY-1998; 98US-085700P.
PR 18-MAY-1998; 98US-086023P.
PR 22-MAY-1998; 98US-086392P.
PR 22-MAY-1998; 98US-086486P.
PR 28-MAY-1998; 98US-087098P.
PR 28-MAY-1998; 98US-087208P.
PR 02-JUN-1998; 98US-087609P.
PR 02-JUN-1998; 98US-087759P.
PR 03-JUN-1998; 98US-087827P.
PR 04-JUN-1998; 98US-088025P.
PR 04-JUN-1998; 98US-088028P.
PR 04-JUN-1998; 98US-088029P.
PR 04-JUN-1998; 98US-088033P.
PR 04-JUN-1998; 98US-088326P.
PR 05-JUN-1998; 98US-088167P.
PR 05-JUN-1998; 98US-088202P.
PR 05-JUN-1998; 98US-088212P.
PR 05-JUN-1998; 98US-088217P.
PR 09-JUN-1998; 98US-088655P.
PR 10-JUN-1998; 98US-088722P.
PR 10-JUN-1998; 98US-088738P.
PR 10-JUN-1998; 98US-088740P.
PR 10-JUN-1998; 98US-088811P.
PR 10-JUN-1998; 98US-088824P.
PR 10-JUN-1998; 98US-088825P.
PR 10-JUN-1998; 98US-088826P.
PR 11-JUN-1998; 98US-088861P.
PR 11-JUN-1998; 98US-088863P.
PR 11-JUN-1998; 98US-088876P.
PR 12-JUN-1998; 98US-089090P.
PR 12-JUN-1998; 98US-089105P.
PR 16-JUN-1998; 98US-089512P.
PR 16-JUN-1998; 98US-089514P.
PR 17-JUN-1998; 98US-089538P.
PR 17-JUN-1998; 98US-089598P.
PR 17-JUN-1998; 98US-089653P.
PR 18-JUN-1998; 98US-089908P.
PR 19-JUN-1998; 98US-089952P.
PR 22-JUN-1998; 98US-090246P.
PR 22-JUN-1998; 98US-090252P.
PR 22-JUN-1998; 98US-090254P.
PR 24-JUN-1998; 98US-090429P.
PR 24-JUN-1998; 98US-090435P.
PR 24-JUN-1998; 98US-090444P.
PR 24-JUN-1998; 98US-090461P.
PR 24-JUN-1998; 98US-090535P.
PR 24-JUN-1998; 98US-090540P.
PR 25-JUN-1998; 98US-090676P.
PR 25-JUN-1998; 98US-090678P.
PR 25-JUN-1998; 98US-090688P.
PR 25-JUN-1998; 98US-090690P.
PR 25-JUN-1998; 98US-090694P.

PR	25-JUN-1998;	98US-090695P.	PR	RESULT 14	
PR	25-JUN-1998;	98US-090696P.	PR	ABU67368	
PR	26-JUN-1998;	98US-090862P.	ID	ABU67368 standard; Protein; 382 AA.	
PR	26-JUN-1998;	98US-090863P.	XX		
PR	26-JUN-1998;	98US-091010P.	AC	ABU67368;	
PR	01-JUL-1998;	98US-091359P.	XX		
PR	01-JUL-1998;	98US-091544P.	DT	29-MAY-2003 (first entry)	
PR	02-JUL-1998;	98US-091486P.	XX		
PR	02-JUL-1998;	98US-091626P.	DE	Human secreted protein PRO234.	
PR	02-JUL-1998;	98US-091628P.	XX		
PR	02-JUL-1998;	98US-091632P.	KW	Human; gene therapy; mucosal lesion; ulcer; enterocolitis; skin disease;	
PR	24-JUL-1998;	98US-094006P.	KW	psoriasis; cancer; lung cancer; colon cancer; nerve cell disease;	
PR	04-AUG-1998;	98US-095282P.	KW	Alzheimer's disease; Parkinson's disease; Usher syndrome; angiogenesis;	
PR	10-AUG-1998;	98US-095998P.	KW	atrophia areata; inflammatory disease; asthma; rheumatoid arthritis;	
PR	10-AUG-1998;	98US-096012P.	XX	ischaemia.	
PR	17-AUG-1998;	98US-096757P.	OS	Homo sapiens.	
PR	17-AUG-1998;	98US-096766P.	XX		
PR	17-AUG-1998;	98US-096867P.	PN	US2003023054-A1.	
PR	17-AUG-1998;	98US-096891P.	XX		
PR	17-AUG-1998;	98US-096897P.	PD	30-JAN-2003.	
PR	18-AUG-1998;	98US-096949P.	XX		
PR	18-AUG-1998;	98US-096959P.	PF	16-JUL-2001; 2001US-0906742.	
PR	18-AUG-1998;	98US-097022P.	XX		
PR	26-AUG-1998;	98US-097952P.	PR	10-SEP-1998; 98WO-US18824.	
PR	26-AUG-1998;	98US-097954P.	PR	14-SEP-1998; 98WO-US19177.	
PR	26-AUG-1998;	98US-097955P.	PR	16-SEP-1998; 98WO-US19330.	
PR	26-AUG-1998;	98US-097971P.	PR	17-SEP-1998; 98WO-US19437.	
PR	26-AUG-1998;	98US-097974P.	PR	01-DEC-1998; 98WO-US25108.	
PR	26-AUG-1998;	98US-098014P.	PR	08-SEP-1999; 99WO-US20594.	
PR	01-SEP-1998;	98US-098716P.	PR	13-SEP-1999; 99WO-US20944.	
PR	01-SEP-1998;	98US-098723P.	PR	15-SEP-1999; 99WO-US21090.	
PR	02-SEP-1998;	98US-098803P.	PR	15-SEP-1999; 99WO-US21547.	
PR	02-SEP-1998;	98US-098821P.	PR	05-OCT-1999; 99WO-US23089.	
PR	02-SEP-1998;	98US-098843P.	PR	29-NOV-1999; 99WO-US28214.	
PR	09-SEP-1998;	98US-099602P.	PR	30-NOV-1999; 99WO-US28313.	
PR	10-SEP-1998;	98US-099741P.	PR	01-DEC-1999; 99WO-US28301.	
PR	10-SEP-1998;	98US-099754P.	PR	02-DEC-1999; 99WO-US28564.	
PR	10-SEP-1998;	98US-099763P.	PR	02-DEC-1999; 99WO-US28565.	
			PR	16-DEC-1999; 99WO-US30095.	
			PR	20-DEC-1999; 99WO-US30911.	
			PR	20-DEC-1999; 99WO-US30999.	
			PR	05-JAN-2000; 2000WO-US00219.	
			PR	11-FEB-2000; 2000WO-US03565.	
			PR	22-FEB-2000; 2000WO-US04414.	
			PR	24-FEB-2000; 2000WO-US05004.	
			PR	02-MAR-2000; 2000WO-US05841.	
			PR	20-MAR-2000; 2000WO-US07377.	
			PR	30-MAR-2000; 2000WO-US08439.	
			PR	22-MAY-2000; 2000WO-US14042.	
			PR	02-JUN-2000; 2000WO-US15264.	
			PR	28-JUL-2000; 2000WO-US20710.	
			PR	24-AUG-2000; 2000WO-US23328.	
			PR	17-SEP-1997; 97US-059113P.	
			PR	17-SEP-1997; 97US-059115P.	
			PR	17-SEP-1997; 97US-059117P.	
			PR	17-SEP-1997; 97US-059119P.	
			PR	17-SEP-1997; 97US-059121P.	
			PR	17-SEP-1997; 97US-059122P.	
			PR	17-SEP-1997; 97US-059184P.	
			PR	18-SEP-1997; 97US-059263P.	
			PR	18-SEP-1997; 97US-059266P.	
			PR	15-OCT-1997; 97US-062125P.	
			PR	17-OCT-1997; 97US-062285P.	
			PR	17-OCT-1997; 97US-062287P.	
			PR	21-OCT-1997; 97US-063486P.	
			PR	24-OCT-1997; 97US-062814P.	
			PR	24-OCT-1997; 97US-062816P.	
			PR	24-OCT-1997; 97US-063045P.	
			PR	24-OCT-1997; 97US-063120P.	
			PR	24-OCT-1997; 97US-063121P.	
			PR	24-OCT-1997; 97US-063127P.	
			PR	24-OCT-1997; 97US-063128P.	

Query Match

Best Local Similarity

Matches 374; Conservative

99.3%;

97.9%;

0;

Score 1986;

Pred. No. 7.2e-183;

Mismatches 0;

DB 24;

Indels 8;

Gaps 1;

Length 382;

QY	1	MRPGTALQAVLLAVLLVGLRAATGRLLS-----GQVCRGGTQRCYKVIYFHDTSR	52
Db	1	MRPGTALQAVLLAVLLVGLRAATGRLLSASDLRLGGQPVCRGGTQRCYKVIYFHDTSR	60
QY	53	RLNFEAKEACRRDGGQLVSIIESEDEQKLEKFIENLLPSDGDFFWIGLRRRREKQSNSTA	112
Db	61	RLNFEAKEACRRDGGQLVSIIESEDEQKLEKFIENLLPSDGDFFWIGLRRRREKQSNSTA	120
QY	113	CQDLYAWTDGSIISQFRNWWYVDEPSCGSEVCVVMYHOPSAAPAGIGGPFYMFQWDDRCNMKN	172
Db	121	CQDLYAWTDGSIISQFRNWWYVDEPSCGSEVCVVMYHOPSAAPAGIGGPFYMFQWDDRCNMKN	180
QY	173	NFICKYSDEKPAVPSREAEGEETELTTPVLPEETOEDAKKTFKESREAAALNLAYILIPS	232
Db	181	NFICKYSDEKPAVPSREAEGEETELTTPVLPEETOEDAKKTFKESREAAALNLAYILIPS	240
QY	233	IPLLLLLVVTTVVCWVWICRKRKREQPDPTKKQHTIWPSPHQNSPDLEVNVIRKQSE	292
Db	241	IPLLLLLVVTTVVCWVWICRKRKREQPDPTKKQHTIWPSPHQNSPDLEVNVIRKQSE	300
QY	293	ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNNAVNPSESGFVTLVSVESGFVTNDIYE	352
Db	301	ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNNAVNPSESGFVTLVSVESGFVTNDIYE	360
QY	353	FSPDQMGRSKESGWVENEIYGY	374
Db	361	FSPDQMGRSKESGWVENEIYGY	382

PR 27-OCT-1997; 97US-063327P.
PR 27-OCT-1997; 97US-063329P.
PR 28-OCT-1997; 97US-063541P.
PR 28-OCT-1997; 97US-063542P.
PR 28-OCT-1997; 97US-063544P.
PR 28-OCT-1997; 97US-063549P.
PR 28-OCT-1997; 97US-063550P.
PR 28-OCT-1997; 97US-063564P.
PR 29-OCT-1997; 97US-063435P.
PR 29-OCT-1997; 97US-063704P.
PR 29-OCT-1997; 97US-063732P.
PR 29-OCT-1997; 97US-063734P.
PR 29-OCT-1997; 97US-063735P.
PR 29-OCT-1997; 97US-063738P.
PR 29-OCT-1997; 97US-064215P.
PR 31-OCT-1997; 97US-063870P.
PR 31-OCT-1997; 97US-064103P.
PR 03-NOV-1997; 97US-064248P.
PR 07-NOV-1997; 97US-064809P.
PR 12-NOV-1997; 97US-065186P.
PR 17-NOV-1997; 97US-065846P.
PR 18-NOV-1997; 97US-065693P.
PR 21-NOV-1997; 97US-066120P.
PR 21-NOV-1997; 97US-066364P.
PR 24-NOV-1997; 97US-066453P.
PR 24-NOV-1997; 97US-066466P.
PR 24-NOV-1997; 97US-066511P.
PR 24-NOV-1997; 97US-066770P.
PR 24-NOV-1997; 97US-066772P.
PR 25-NOV-1997; 97US-066840P.
PR 12-DEC-1997; 97US-069425P.
PR 04-JUN-1998; 98US-088026P.
PR 10-SEP-1998; 98US-099803P.
PR 14-SEP-1998; 98US-100262P.
PR 17-SEP-1998; 98US-100858P.
PR 13-OCT-1998; 98US-104080P.
PR 20-NOV-1998; 98US-109304P.
PR 22-DEC-1998; 98US-113296P.
PR 07-JUL-1999; 99US-143048P.
PR 26-JUL-1999; 99US-145698P.
PR 28-JUL-1999; 99US-146222P.
PR 18-SEP-2000; 2000US-0665350.
XX
PA (GETH) GENENTECH INC.
XX
PI Ashkenazi A, Botstein D, Desnoyers L, Eaton DL, Ferrara N;
PI Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME, Goddard A;
PI Godowski PJ, Grimaldi JC, Gurney AL, Hillan KJ, Kljavin IJ;
PI Mather JP, Pan J, Paoni NF, Roy MA, Stewart TA, Tumas D;
PI Williams PM, Wood WI;
XX
DR WPI; 2003-331485/31.
DR N-PSDB; ACA05450.
XX
PT Sixty one isolated nucleic acids encoding a PRO polypeptide, e.g.
PT PRO245 or PRO1868, useful in chromosome and gene mapping, in generating
PT antisense RNA and DNA, and in treating cancer and Alzheimer's disease -
XX
PS Example 22; Fig 50; 481pp; English.
XX
CC The invention relates to sixty one nucleic acids encoding PRO
CC polypeptides (secreted and transmembrane). The polynucleotide is useful
CC in molecular biology, including uses as hybridisation probes, in
CC chromosome and gene mapping, in generating antisense RNA and DNA, and in
CC gene therapy. The polynucleotide may also be used in preparing PRO
CC polypeptides by recombinant techniques, and in generating either
CC transgenic animals or knock-out animals which, in turn, are useful in the
CC development and screening of therapeutically useful reagents. The PRO
CC polypeptide or the antibody is used in preparing a medicament for
CC treating a condition responsive to the polypeptide or antibody, such as
CC mucosal lesions e.g. ulcers and enterocolitis, skin disease e.g.
CC psoriasis, cancer e.g. lung cancer and colon cancer, nerve cell disease
CC e.g. Alzheimer's disease and Parkinson's disease, Usher syndrome,

CC atrophla areata, angiogenesis, inflammation, inflammatory disease e.g asthma and
CC rheumatoid arthritis, ischaemia, and in various diagnostic assays. The
CC present sequence represents the amino acid sequence of a PRO polypeptide.
XX
SQ Sequence 382 AA;
Query Match 99.3%; Score 1986; DB 24; Length 382;
Best Local Similarity 97.9%; Pred. No. 7.2e-183;
Matches 374; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
QY 1 MRPGTALQAVLLAVLLVGLRAATGRLLS-----GQVCRGGTQRPCYKVIYFHDTSR 52
|||
DB 1 MRPGTALQAVLLAVLLVGLRAATGRLLSASDLDRGGQPVCRGGTQRPCYKVIYFHDTSR 60
|||
QY 53 RLFEEAKEACRRDGGQLVSISEDEQKLEKPIENLLPSDGDGFWIGLRRREEKQSNSTA 112
|||
DB 61 RLFEEAKEACRRDGGQLVSISEDEQKLEKPIENLLPSDGDGFWIGLRRREEKQSNSTA 120
|||
QY 113 CQDLYAWTDGSIQFRNWIYVDEPSCGSEVCVVMYHQPSAPAGIGGYPYMFQWDDRCNMKN 172
|||
DB 121 CQDLYAWTDGSIQFRNWIYVDEPSCGSEVCVVMYHQPSAPAGIGGYPYMFQWDDRCNMKN 180
|||
QY 173 NFICKYSDEKPAVPSREAEGETELTTPVLPEETQEEADAKTKFKESREALNLAYILIPS 232
|||
DB 181 NFICKYSDEKPAVPSREAEGETELTTPVLPEETQEEADAKTKFKESREALNLAYILIPS 240
|||
QY 233 IPLLALLVVTTCVWVICRKRKREQPDSTKQHTIWPSPHQGNSPDLEVYNVIRKQSE 292
|||
DB 241 IPLLALLVVTTCVWVICRKRKREQPDSTKQHTIWPSPHQGNSPDLEVYNVIRKQSE 300
|||
QY 293 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTILSVESGFVTNDIYE 352
|||
DB 301 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTILSVESGFVTNDIYE 360
|||
QY 353 FSPDQMGRSKESGWVENEIYGY 374
|||
DB 361 FSPDQMGRSKESGWVENEIYGY 382
|||
RESULT 15
ABU67415
ID ABU67415 standard; Protein; 382 AA.
XX
AC ABU67415;
XX
DT 29-MAY-2003 (first entry)
XX
DE Human secreted/transmembrane protein (PRO) #10.
XX
KW Human; secreted and transmembrane protein; PRO; TNF-alpha;
KW tumour necrosis factor alpha; chondrocyte cell; tumour; gene therapy;
KW tissue typing.
XX
OS Homo sapiens.
XX
PN US2003036162-A1.
XX
PD 20-FEB-2003.
XX
PF 12-JUL-2002; 2002US-0194423.
XX
PR 16-SEP-1998; 98WO-US19330.
PR 07-OCT-1998; 98WO-US21141.
PR 01-DEC-1998; 98WO-US25108.
PR 08-MAR-1999; 99WO-US05028.
PR 14-MAY-1999; 99WO-US10733.
PR 02-JUN-1999; 99WO-US12252.
PR 01-SEP-1999; 99WO-US20111.
PR 15-SEP-1999; 99WO-US21090.
PR 01-DEC-1999; 99WO-US28301.
PR 02-DEC-1999; 99WO-US28551.
PR 30-DEC-1999; 99WO-US31274.
PR 05-JAN-2000; 2000WO-US00219.

PR 18-FEB-2000; 2000WO-US04341.
PR 18-FEB-2000; 2000WO-US04342.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US05004.
PR 01-MAR-2000; 2000WO-US05601.
PR 02-MAR-2000; 2000WO-US05841.
PR 15-MAR-2000; 2000WO-US06884.
PR 30-MAR-2000; 2000WO-US08439.
PR 17-MAY-2000; 2000WO-US13705.
PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.
PR 02-JUN-2000; 2000WO-US15264.
PR 28-JUL-2000; 2000WO-US20710.
PR 24-AUG-2000; 2000WO-US23328.
PR 08-NOV-2000; 2000WO-US30952.
PR 01-DEC-2000; 2000WO-US32678.
PR 20-DEC-2000; 2000WO-US34956.
PR 28-FEB-2001; 2001WO-US06520.
PR 01-JUN-2001; 2001WO-US17800.
PR 20-JUN-2001; 2001WO-US19692.
PR 29-JUN-2001; 2001WO-US21066.
PR 09-JUL-2001; 2001WO-US21735.
PR 29-AUG-2001; 2001WO-US27099.
PR 26-JUN-1998; 98US-0105413.
PR 07-OCT-1998; 98US-0168978.
PR 06-NOV-1998; 98US-0187368.
PR 07-DEC-1998; 98US-0202054.
PR 03-MAR-1999; 99US-0254311.
PR 14-MAY-1999; 99US-0311832.
PR 14-MAY-1999; 99US-0380137.
PR 25-AUG-1999; 99US-0380138.
PR 25-AUG-1999; 99US-0380139.
PR 25-AUG-1999; 99US-0380142.
PR 18-OCT-1999; 99US-0403297.
PR 12-NOV-1999; 99US-0423844.
PR 22-AUG-2000; 2000US-0644848.
PR 18-SEP-2000; 2000US-0664610.
PR 18-SEP-2000; 2000US-0665350.
PR 08-NOV-2000; 2000US-0709238.
PR 20-DEC-2000; 2000US-0747259.
PR 22-MAR-2001; 2001US-0816744.
PR 10-MAY-2001; 2001US-0854208.
PR 10-MAY-2001; 2001US-0854280.
PR 25-MAY-2001; 2001US-0866028.
PR 05-JUN-2001; 2001US-0874503.
PR 18-JUL-2001; 2001US-0908827.
PR 30-JUL-2001; 2001US-0918585.
PR 06-AUG-2001; 2001US-0924419.
PR 13-AUG-2001; 2001US-0929404.
PR 16-AUG-2001; 2001US-0931836.
PR 28-AUG-2001; 2001US-0941992.
PR 04-SEP-2001; 2001US-0946374.
PR 15-JAN-2002; 2002US-0052586.

(GETH) GENENTECH INC.

Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;
Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;

WPI; 2003-332039/31.
N-PSDB; ACA05709.

New secreted and transmembrane PRO polypeptides and nucleic acids,
useful in gene therapy, in chromosome and gene mapping, as chromosome
markers, in tissue typing, and in chromosome identification -

Claim 11; Fig 20; 706pp; English.

The invention discloses human nucleic acids encoding secreted and
transmembrane (PRO) polypeptides. Also disclosed is an antibody that
specifically binds to the PRO polypeptide, a method for stimulating the
release of tumour necrosis factor alpha (TNF-alpha) from human blood by
contacting the blood a PRO polypeptide, a method for stimulating the

CC proliferation or differentiation of chondrocyte cells by contacting the
CC cells with a PRO polypeptide, a method for detecting the presence of a
CC tumour in a mammal and an oligonucleotide probe derived from any of the
CC PRO nucleotide sequences. The nucleotide sequences are useful as probes,
CC in chromosome and gene mapping, in generating antisense RNA and DNA, in
CC preparing PRO polypeptides by recombinant techniques and in gene therapy
CC (e.g. for replacement of defective gene). The PRO polypeptides are useful
CC as molecular weight markers for protein electrophoresis purposes, for
CC chromosome identification, as chromosome markers, as therapeutic agents,
CC for stimulating the release of TNF-alpha from human blood, for
CC stimulating the proliferation or differentiation of chondrocytes and
CC detecting the presence of a tumour. The PRO polypeptides and nucleic
CC acids may also be used diagnostically for tissue typing. The sequences
CC presented in ABU67406-ABU67710 are the PRO polypeptides of the invention.
XX
SQ Sequence 382 AA;

Query Match 99.3%; Score 1986; DB 24; Length 382;
Best Local Similarity 97.9%; Pred. No. 7.2e-183;
Matches 374; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
QY 1 MRPGTALQAVLLAVLLVGLRAATGRLLS-----GQVCRGGTQPCYKVIYFHDTSR 52
Db 1 MRPGTALQAVLLAVLLVGLRAATGRLLSASDLDLRGGQPCVCRGGTQPCYKVIYFHDTSR 60
QY 53 RLNFEEAKEACRRDGGQLVSIESEDEQKLIIEKFIEIENLLPSDGFWIGLRRREEKQSNSTA 112
Db 61 RLNFEEAKEACRRDGGQLVSIESEDEQKLIIEKFIEIENLLPSDGFWIGLRRREEKQSNSTA 120
QY 113 CQDLYAWTDGSIQFRNWWYVDEPSCGSEVCVVMYHQPSAPAGIGGPFQWDDRCNMKN 172
Db 121 CQDLYAWTDGSIQFRNWWYVDEPSCGSEVCVVMYHQPSAPAGIGGPFQWDDRCNMKN 180
QY 173 NFICKYSDEKPAVPSREAEGETELTTPVLPEETQEEEDAKKTFKESREAAALNLAYILIPS 232
Db 181 NFICKYSDEKPAVPSREAEGETELTTPVLPEETQEEEDAKKTFKESREAAALNLAYILIPS 240
QY 233 IPLLLLLVVTWVCWWVICRKRKEQDPSTKKQHTIWPSPHQGNSPDLEVNVIRKQSE 292
Db 241 IPLLLLLVVTWVCWWVICRKRKEQDPSTKKQHTIWPSPHQGNSPDLEVNVIRKQSE 300
QY 293 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFTNDIYE 352
Db 301 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFTNDIYE 360
QY 353 FSPDQGRSKESGWVENEIYGY 374
Db 361 FSPDQGRSKESGWVENEIYGY 382

Search completed: December 22, 2003, 16:10:06
Job time : 54.9807 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 22, 2003, 16:04:13 ; Search time 28.8986 Seconds
(without alignments)
1120.478 Million cell updates/sec

Title: US-09-887-855-2_COPY_24_227

Perfect score: 1106

Sequence: 1 GRLLSGQPVCRCGGTQPCVK.....EEDAKKTFKESREAAALNLAY 204

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_19Jun03.*
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1106	100.0	374	21 AAY93948	Amino acid sequenc
2	1106	100.0	374	22 AAE03651	Human extracellular
3	1106	100.0	374	23 ABG66680	Human novel polype
4	1106	100.0	374	23 ABB90203	Human polypeptide
5	1106	100.0	387	22 AAM25796	Human protein sequ
6	1096	99.1	374	21 AAY91490	Human secreted pro
7	1092	98.7	382	20 AAY13367	Amino acid sequenc
8	1092	98.7	382	22 AAU29033	Human PRO polypept
9	1092	98.7	382	22 AAB80235	Human PRO234 prote

10	1092	98.7	382	24	ABU69645	Novel human secret
11	1092	98.7	382	24	ABU71121	Human PRO234 prote
12	1092	98.7	382	24	ABU71468	Human PRO polypept
13	1092	98.7	382	24	ABU71914	Human secreted/tra
14	1092	98.7	382	24	ABU65578	Human secreted/tra
15	1092	98.7	382	24	ABU65911	Novel human secret
16	1092	98.7	382	24	ABU67368	Human secreted pro
17	1092	98.7	382	24	ABU67415	Human secreted/tra
18	1092	98.7	382	24	ABU64522	Human secreted/tra
19	1092	98.7	382	24	ABU65273	Human PRO polypept
20	1092	98.7	382	24	ABU58409	Human PRO polypept
21	1092	98.7	382	24	ABU55945	Human secreted/tra
22	1092	98.7	382	24	ABU56940	Human PRO polypept
23	1092	98.7	382	24	ABU54370	Human secreted/tra
24	1092	98.7	382	24	ABU10519	Human secreted/tra
25	561.5	50.8	273	21	AAB18913	A novel polypeptid
26	561.5	50.8	273	22	AAU12441	Human PRO1890 poly
27	561.5	50.8	273	22	AAB73309	Human C-type lecti
28	561.5	50.8	273	22	AAB87609	Human PRO1890. Ho
29	561.5	50.8	273	23	ABG95934	Human secreted/tra
30	561.5	50.8	273	23	ABG95554	Human angiogenesis
31	561.5	50.8	273	23	ABB84948	Human PRO1890 prot
32	561.5	50.8	273	24	ABU69084	Human PRO polypept
33	561.5	50.8	273	24	ABU69107	Human PRO polypept
34	561.5	50.8	273	24	ABU71589	Human secreted pol
35	561.5	50.8	273	24	ABU72035	Novel human secret
36	561.5	50.8	273	24	ABU72192	Human PRO polypept
37	561.5	50.8	273	24	ABU66839	Human PRO polypept
38	561.5	50.8	273	24	ABU67115	Human secreted/tra
39	561.5	50.8	273	24	ABU59920	Novel secreted and
40	534.5	48.3	274	23	ABP69211	Human polypeptide
41	504.5	45.6	232	22	AAB94192	Human protein sequ
42	275	24.9	102	22	ABG51278	Human liver peptid
43	275	24.9	102	22	ABB21788	Protein #3787 enco
44	275	24.9	102	22	AAU29941	Peptide #3978 enco
45	231	20.9	81	21	AAY91643	Human secreted pro

ALIGNMENTS

RESULT 1

AAAY93948

ID AAY93948 standard; Protein; 374 AA.

XX

AC AAY93948;

XX

DT 03-OCT-2000 (first entry)

XX

DE Amino acid sequence of a lectin ss3939 polypeptide.

XX

KW Human; lectin ss3939; chromosome 11; gene therapy.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT Peptide 1..21

FT Domain /note= "signal peptide"

FT Domain 22..227

FT Domain /note= "extracellular coding region"

FT Domain 228..248

FT Domain /note= "predicted transmembrane domain"

FT Domain 249..374

FT Domain /note= "predicted cytoplasmic or intracellular domain"

XX

PN WO200039296-A1.

XX

PD 06-JUL-2000.

XX

PF 22-DEC-1999; 99WO-US30523.

XX

PR 23-DEC-1998; 98US-0113820.

XX

QY 1 GRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI ESEDEQKLI E 60
Db 24 GRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI ESEDEQKLI E 83
QY 61 KFIENLLPSDGD F WIGLRRREEKQSNSTACQDLYAWTDGSI SQFRN WYVDEPSCGSEVCV 120
Db 84 KFIENLLPSDGD F WIGLRRREEKQSNSTACQDLYAWTDGSI SQFRN WYVDEPSCGSEVCV 143
QY 121 VMYHQPSAPAGIGGPMYFQWNDRCNMKNPFICKYSDEKPAVPSREAEGETELTTPVLP 180
Db 144 VMYHQPSAPAGIGGPMYFQWNDRCNMKNPFICKYSDEKPAVPSREAEGETELTTPVLP 203
QY 181 EETQEEDAKKTFKESREAAALNLAY 204
Db 204 EETQEEDAKKTFKESREAAALNLAY 227

RESULT 3

ABG66680

ID ABG66680 standard; Protein; 374 AA.

XX

AC ABG66680;

XX

DT 30-AUG-2002 (first entry)

XX

DE Human novel polypeptide #15.

XX

KW Human; inflammatory condition; shock; sepsis; immune response;

KW cancer; wound healing; central nervous system disease; haematopoiesis;

KW peripheral nervous system disease; amyotrophic lateral sclerosis; tendon;

KW myeloid cell disorder; lymphoid cell disorder; platelet disorder; bone;

KW cartilage; ligament; nerve tissue; ulcer; osteoporosis; osteoarthritis;

KW bone degenerative disorder; periodontal disease; reperfusion injury;

KW lung fibrosis; liver fibrosis; autoimmune disorder; bacterial infection;

KW allergic condition; thrombolysis; thrombosis; coagulation disorder;

KW fungal infection.

XX

OS Homo sapiens.

XX

PN WO200244340-A2.

XX

PD 06-JUN-2002.

XX

PF 30-NOV-2001; 2001WO-US47004.

XX

PR 30-NOV-2000; 2000US-0028952.

XX

PA (HYSE-) HYSEQ INC.

XX

PI Tang YT, Goodrich RW, Liu C, Zhou P, Asundi V, Wang J, Wang D;

PI Yamazaki V, Ujwal ML, Drmanac RT;

XX

DR WPI; 2002-508509/54.

DR N-PSDB; ABK94904.

XX

PT Novel nucleic acids and polypeptides for diagnosis, treatment of

PT inflammatory, autoimmune, nervous system, myeloid or lymphoid cell

PT disorders, cancer and promoting wound healing -

XX

PS Claim 10; Page 579-580; 672pp; English.

XX

CC The invention relates to human novel polynucleotides and associated

CC polypeptides. The polynucleotides and polypeptides are useful for

CC treating inflammatory conditions such as arthritis, nephritis, Crohn's

CC disease, ischaemia-reperfusion injury, shock, sepsis, immune responses

CC and cancer and for promoting wound healing. The sequences are used to

CC induce the proliferation of neural cells and regeneration of nerve and

CC brain tissue, and are useful for the treatment of central and peripheral

CC nervous system diseases and neuropathies, such as Alzheimer's disease,

CC Parkinson's disease, Huntington's disease and amyotrophic lateral

CC sclerosis. The sequences are involved in chemotactic or chemokinetic

CC activity, regulation of haematopoiesis, treatment of myeloid or lymphoid

CC cell disorders and platelet disorders such as thrombocytopenia,
CC regeneration of bone, cartilage, tendon, ligament and/or nerve tissue
CC growth, tissue repair, healing of burns, incisions, ulcers, treatment of
CC osteoporosis, osteoarthritis, bone degenerative disorders and periodontal
CC disease. The sequences of the invention are also useful for gut
CC protection or regeneration and treatment of lung or liver fibrosis,
CC reperfusion injury in various tissues, immune deficiencies and disorders
CC including severe combined immunodeficiency (SCID), bacterial or fungal
CC infections, autoimmune disorders e.g. multiple sclerosis and myasthenia
CC gravis, allergic conditions such as asthma, thrombolysis or thrombosis
CC and coagulation disorders. Sequences ABG66666-ABG66758 represent human
CC novel polypeptides of the invention.

XX

SQ Sequence 374 AA;

Query Match 100.0%; Score 1106; DB 23; Length 374;

Best Local Similarity 100.0%; Pred. No. 2.5e-104;

Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI ESEDEQKLI E 60

Db 24 GRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI ESEDEQKLI E 83

QY 61 KFIENLLPSDGD F WIGLRRREEKQSNSTACQDLYAWTDGSI SQFRN WYVDEPSCGSEVCV 120

Db 84 KFIENLLPSDGD F WIGLRRREEKQSNSTACQDLYAWTDGSI SQFRN WYVDEPSCGSEVCV 143

QY 121 VMYHQPSAPAGIGGPMYFQWNDRCNMKNPFICKYSDEKPAVPSREAEGETELTTPVLP 180

Db 144 VMYHQPSAPAGIGGPMYFQWNDRCNMKNPFICKYSDEKPAVPSREAEGETELTTPVLP 203

QY 181 EETQEEDAKKTFKESREAAALNLAY 204

Db 204 EETQEEDAKKTFKESREAAALNLAY 227

RESULT 4

ABB90203

ID ABB90203 standard; Protein; 374 AA.

XX

AC ABB90203;

XX

DT 24-MAY-2002 (first entry)

XX

DE Human polypeptide SEQ ID NO 2579.

XX

KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;

KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;

KW vulnery; anticonvulsant; antibacterial; antifungal; antiparasitic;

KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;

KW neurological disease; infection; human; secreted protein.

XX

OS Homo sapiens.

XX

PN WO200190304-A2.

XX

PD 29-NOV-2001.

XX

PF 18-MAY-2001; 2001WO-US16450.

XX

PR 19-MAY-2000; 2000US-205515P.

XX

PA (HUMA-) HUMAN GENOME SCI INC.

XX

PI Birse CE, Rosen CA;

XX

DR WPI; 2002-122018/16.

DR N-PSDB; ABL90612.

XX

PT Novel 1405 isolated polypeptides, useful for diagnosis, treatment and

PT prevention of neural, immune system, muscular, reproductive,

PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative

PT disorders -

PS Claim 11; SEQ ID NO 2579; 2081pp + Sequence Listing; English.

XX The invention relates to novel genes (ABL9449-ABL90853) and proteins

CC (ABB9040-ABB9044) useful for preventing, treating or ameliorating

CC medical conditions e.g. by protein or gene therapy. The genes are

CC isolated from a range of human tissues disclosed in the specification.

CC The nucleic acids, proteins, antibodies and (ant)agonists are useful

CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast

CC and ovarian cancer and other cancers of the adrenal gland, bone, bone

CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;

CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune

CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's

CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative

CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;

CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and

CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal

CC and parasitic infections.

CC Note: The sequence data for this patent did not form part of the

CC printed specification, but was obtained in electronic format directly

CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX

SQ Sequence 374 AA;

Query Match 100.0%; Score 1106; DB 23; Length 374;

Best Local Similarity 100.0%; Pred. No. 2.5e-104;

Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRLLSGQPVCRGGTQPCYKVIYFHDTSRRLNFEAKEACRRDGGQLVSISEDEQKLIIE 60

Db 24 GRLLSGQPVCRGGTQPCYKVIYFHDTSRRLNFEAKEACRRDGGQLVSISEDEQKLIIE 83

QY 61 KFIENLLPSDGDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNWWYVDEPSCGSEVCV 120

Db 84 KFIENLLPSDGDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNWWYVDEPSCGSEVCV 143

QY 121 VMYHQPSAPAGIGGPFYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGETELTTPVLP 180

Db 144 VMYHQPSAPAGIGGPFYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGETELTTPVLP 203

QY 181 EETQEEDAKKTFKESREAAALNLAY 204

Db 204 EETQEEDAKKTFKESREAAALNLAY 227

RESULT 5

AAM25796

ID AAM25796 standard; Protein; 387 AA.

XX

AC AAM25796;

XX

DT 16-OCT-2001 (first entry)

XX

DE Human protein sequence SEQ ID NO:1311.

XX

KW Human; cancer; ulcer; HIV infection; human immunodeficiency virus;

KW antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;

KW antibacterial; endocrine; cardiant; central nervous system; virucide;

KW anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;

KW antiaggregant; haemostatic; vulnery; antiulcer; osteopathic; eczema;

KW dermatological; antiallergic; antiasthmatic; antidiabetic; cytostatic;

KW neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;

KW immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;

KW antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;

KW cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;

KW genetic disease; haematopoietic disorder; platelet disorder; asthma;

KW thrombocytopaenia; osteoporosis; severe combined immunodeficiency;

KW allergic rhinitis; diabetes; multiple sclerosis; depression;

KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;

XX

OS Homo sapiens.

XX

PN WO200153455-A2.

XX

PD 26-JUL-2001.

XX

PF 22-DEC-2000; 2000WO-US35017.

XX

PR 23-DEC-1999; 99US-0471275.

PR 21-JAN-2000; 2000US-0488725.

PR 25-APR-2000; 2000US-0552317.

XX

PA (HYSE-) HYSEQ INC.

XX

PI Tang YT, Liu C, Drmanac RT;

XX

DR WPI; 2001-457603/49.

DR N-PSDB; AAH99737.

XX

PT Isolated human polynucleotides encoding polypeptides, useful for the

PT treatment and diagnosis of e.g. cancer, ulcers and HIV infection -

XX

PS Claim 20; Page 272; 1217pp; English.

XX

CC AAH99166 to AAH99904 encode the human proteins given in AAM25225 to

CC AAM25963. The proteins can have activities based on the tissues and

CC cells they are expressed in, such as: antiinflammatory; antirheumatic;

CC antiarthritic; immunosuppressive; antibacterial; endocrine; cardiant;

CC central nervous system; virucide; anti-HIV; fungicide; antimutagen;

CC cardiovascular; antianaemic; antiaggregant; haemostatic; vulnery;

CC antiulcer; osteopathic; dermatological; antiallergic; antiasthmatic;

CC antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;

CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides

CC encoding them can be used in gene therapy, antisense therapy and vaccine

CC production. The proteins and polynucleotides are useful for screening for

CC agonists or antagonists of a protein and for the treatment and diagnosis

CC of disorders associated with the activity of a protein e.g. inflammation,

CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,

CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal

CC infections, autoimmunity, genetic diseases, haematopoietic disorders,

CC anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,

CC osteoporosis, severe combined immunodeficiency, eczema, allergic

CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,

CC Alzheimer's disease, Parkinson's disease, neurodegenerative and

CC neurological disorders.

XX

SQ Sequence 387 AA;

Query Match 100.0%; Score 1106; DB 22; Length 387;

Best Local Similarity 100.0%; Pred. No. 2.6e-104;

Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GRLLSGQPVCRGGTQPCYKVIYFHDTSRRLNFEAKEACRRDGGQLVSISEDEQKLIIE 60

Db 37 GRLLSGQPVCRGGTQPCYKVIYFHDTSRRLNFEAKEACRRDGGQLVSISEDEQKLIIE 96

QY 61 KFIENLLPSDGDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNWWYVDEPSCGSEVCV 120

Db 97 KFIENLLPSDGDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNWWYVDEPSCGSEVCV 156

QY 121 VMYHQPSAPAGIGGPFYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGETELTTPVLP 180

Db 157 VMYHQPSAPAGIGGPFYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGETELTTPVLP 216

QY 181 EETQEEDAKKTFKESREAAALNLAY 204

Db 217 EETQEEDAKKTFKESREAAALNLAY 240

RESULT 6

AA91490

ID AA91490 standard; Protein; 374 AA.

XX

AC AA91490;

XX

DT 29-JUN-2000 (first entry)
XX Human secreted protein sequence encoded by gene 40 SEQ ID NO:163.
DE
XX
KW Human; secreted protein; diagnosis; cytostatic; immunosuppressive;
KW antiHIV; antiinflammatory; nootropic; neuroprotective; antiallergic;
KW osteopathic; antiarthritic; antibacterial; antidiabetic; antiasthma;
KW antipsoriatic; cardiant; gene therapy; cancer; neurological disorder;
KW immune disease; inflammation; blood disorder; tumour.
XX
OS Homo sapiens.
XX
XX WO200006698-A1.
XX
XX 10-FEB-2000.
XX
XX 29-JUL-1999; 99WO-US17130.
XX
XX 30-JUL-1998; 98US-0094657.
XX 05-AUG-1998; 98US-0095486.
XX 06-AUG-1998; 98US-0095454.
XX 06-AUG-1998; 98US-0095455.
XX 12-AUG-1998; 98US-0096319.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Komatsoulis GA, Rosen CA, Ruben SM, Duan R, Moore PA, Shi Y;
XX Lafleur D, Wei Y, Ni J, Florence KA, Young PE, Brewer LA;
XX Soppet DR, Endress GA, Ebner R, Olsen HS, Mucenski M;
XX
XX WPI; 2000-195282/17.
XX N-PSDB; AAA26385.
XX
XX New isolated human genes and the secreted polypeptides they encode,
XX useful for diagnosis and treatment of e.g. cancers, neurological
XX disorders, immune diseases, inflammation or blood disorders -
XX
XX Claim 11; Page 483-484; 634pp; English.
XX
XX The polynucleotide sequences given in AAA26346 to AAA26458 encode the
XX human secreted proteins given in AAY91451 to AAY91691. The human secreted
XX proteins can have activities based on the tissues and cells they are
XX expressed in. Examples of the activities are: cytostatic;
XX immunosuppressive; antiHIV; antiinflammatory; nootropic; neuroprotective;
XX antiallergic; osteopathic; antiarthritic; antibacterial; antidiabetic;
XX antiasthma; antipsoriatic; and cardiant. The polynucleotides and their
XX corresponding secreted proteins are useful for preventing, treating or
XX ameliorating medical conditions, e.g. by protein or gene therapy. Also
XX pathological conditions can be diagnosed by determining the amount of the
XX proteins in a sample or by determining the presence of mutations in the
XX polynucleotides. Specific uses are described for each of the
XX polynucleotides, based on which tissues they are most highly expressed
XX in, and include developing products for the diagnosis or treatment of
XX cancer, tumours, neurodegenerative disorders, developmental abnormalities
XX and foetal deficiencies, blood disorders, diseases of the immune system,
XX autoimmune diseases, hepatic and renal disease, inflammation,
XX allergies, Alzheimer's and behavioural disorders, schizophrenia,
XX osteoporosis, arthritis, infections, AIDS, spinal cord injuries,
XX transplant rejection, diabetes, asthma, sepsis, acne, psoriasis,
XX cardiovascular disorders, reproductive disorders, gastrointestinal
XX disorders, respiratory disorders and metabolic disorders. The proteins
XX or polynucleotides can also be used as food additives or preservatives.
XX The proteins are also useful for identifying their binding partners.
XX AAA26337 to AAA26345 and AAY91450 are sequences used in the
XX exemplification of the present invention.
XX
XX Sequence 374 AA;
XX
XX Query Match 99.1%; Score 1096; DB 21; Length 374;
XX Best Local Similarity 99.0%; Pred. No. 2.6e-103;
XX Matches 202; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX 1 GRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEAKEACRRDGGQLVSEDEQKLE 60

Db 24 GRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEAKEACRRDGGQLVSEDEQKLE 83
Qy 61 KFIENLLPSDGDGFWIGLRRREEKQSNSTACODLYAWTDGTSIQFRNMYVDEPSCGSEVCV 120
Db 84 XFIENLLPSDGDGFWIGLRRREEKQSNSTXCQDLYAWTDGTSIQFRNMYVDEPSCGSEVCV 143
Qy 121 VMYHOPSAPAGIGGPMFQWNDRCNMKNFICKYSDEKPAVPSREAEGETELTTPVLP 180
Db 144 VMYHOPSAPAGIGGPMFQWNDRCNMKNFICKYSDEKPAVPSREAEGETELTTPVLP 203
Qy 181 EETQEDAKKTFKESREAAALNLAY 204
Db 204 EETQEDAKKTFKESREAAALNLAY 227
RESULT 7
AAY13367
ID AAY13367 standard; Protein; 382 AA.
XX
AC AAY13367;
XX
DT 25-JUN-1999 (first entry)
XX
DE Amino acid sequence of protein PRO234.
XX
KW Secreted protein; transmembrane protein; human; enterocolitis;
KW Zollinger-Ellison syndrome; gastrointestinal ulceration;
KW congenital microvillus atrophy; skin disease; cell growth;
KW abnormal keratinocyte differentiation; psoriasis; epithelial cancer;
KW Parkinson's disease; Alzheimer's disease; ALS; neuropathy;
KW fibromodulin; dermal scarring; Usher Syndrome; Atrophia areata;
KW anti-thrombotic; wound healing; tissue repair.
XX
OS Homo sapiens.
XX
XX WO9914328-A2.
XX
PD 25-MAR-1999.
XX
XX 16-SEP-1998; 98WO-US19330.
XX
XX 25-NOV-1997; 97US-0066840.
XX 17-SEP-1997; 97US-0059113.
XX 17-SEP-1997; 97US-0059115.
XX 17-SEP-1997; 97US-0059117.
XX 17-SEP-1997; 97US-0059119.
XX 17-SEP-1997; 97US-0059121.
XX 17-SEP-1997; 97US-0059122.
XX 17-SEP-1997; 97US-0059184.
XX 18-SEP-1997; 97US-0059263.
XX 18-SEP-1997; 97US-0059266.
XX 15-OCT-1997; 97US-0062125.
XX 17-OCT-1997; 97US-0062285.
XX 17-OCT-1997; 97US-0062287.
XX 21-OCT-1997; 97US-0063486.
XX 24-OCT-1997; 97US-0062814.
XX 24-OCT-1997; 97US-0062816.
XX 24-OCT-1997; 97US-0063045.
XX 24-OCT-1997; 97US-0063120.
XX 24-OCT-1997; 97US-0063121.
XX 24-OCT-1997; 97US-0063127.
XX 24-OCT-1997; 97US-0063128.
XX 27-OCT-1997; 97US-0063329.
XX 27-OCT-1997; 97US-0063327.
XX 28-OCT-1997; 97US-0063541.
XX 28-OCT-1997; 97US-0063542.
XX 28-OCT-1997; 97US-0063544.
XX 28-OCT-1997; 97US-0063549.
XX 28-OCT-1997; 97US-0063550.
XX 28-OCT-1997; 97US-0063564.
XX 29-OCT-1997; 97US-0063435.
XX 29-OCT-1997; 97US-0063704.

PR 29-OCT-1997; 97US-0063732.
PR 29-OCT-1997; 97US-0063738.
PR 29-OCT-1997; 97US-0063734.
PR 29-OCT-1997; 97US-0064215.
PR 29-OCT-1997; 97US-0063735.
PR 31-OCT-1997; 97US-0063870.
PR 31-OCT-1997; 97US-0064103.
PR 03-NOV-1997; 97US-0064248.
PR 07-NOV-1997; 97US-0064809.
PR 12-NOV-1997; 97US-0065186.
PR 17-NOV-1997; 97US-0065846.
PR 18-NOV-1997; 97US-0065693.
PR 21-NOV-1997; 97US-0066120.
PR 21-NOV-1997; 97US-0066364.
PR 24-NOV-1997; 97US-0066772.
PR 24-NOV-1997; 97US-0066466.
PR 24-NOV-1997; 97US-0066770.
PR 24-NOV-1997; 97US-0066511.
PR 24-NOV-1997; 97US-0066453.
XX
PA (GETH) GENENTECH INC.
XX
PI Chen J, Goddard A, Gurney AL, Pennica D, Wood WI, Yuan J;
XX WPI; 1999-229533/19.
DR N-PSDB; AAX52238.
XX
PT New isolated human genes and polypeptides used in, e.g. treatment of
PT gastrointestinal ulceration
XX
PS Claim 12; Fig 50; 320pp; English.
XX
CC AAY13344-403 represent secreted and transmembrane human proteins.
CC The cDNA sequences are obtained from cDNA libraries, prepared from
CC fetal lung, fetal kidney, fetal brain, fetal liver and fetal retina.
CC The encoded polypeptides have specific uses based on their homology to
CC known polypeptides, e.g. PRO211 and PRO217 can be used for disorders
CC associated with the preservation and maintenance of gastrointestinal
CC mucosa and the repair of acute and chronic mucosal lesions
CC (e.g. enterocolitis, Zollinger-Ellison syndrome, gastrointestinal
CC ulceration and congenital microvillus atrophy), skin diseases associated
CC with abnormal keratinocyte differentiation (e.g. psoriasis, epithelial
CC cancers such as lung squamous cell carcinoma of the vulva and gliomas),
CC potent effects on cell growth and development, diseases related to
CC growth or survival of nerve cells including Parkinson's disease,
CC Alzheimer's disease, ALS, neuropathies or cancer. PRO265 can be used as
CC for fibromodulin, e.g. for reducing dermal scarring. PRO264 can be used
CC as a target for anti-tumor drugs. PRO533 may be used in the treatment
CC of Usher Syndrome or Atrophia areata; PRO269 can be used as an
CC anti-thrombotic agent; PRO287 polypeptides and portions may have
CC therapeutic applications in wound healing and tissue repair; PRO317 can
CC be used for treating problems of the kidney, uterus, endometrium, blood
CC vessels, or related tissue, e.g. in the heart of genital tract.
XX
SQ Sequence 382 AA;

Query Match 98.7%; Score 1092; DB 20; Length 382;
Best Local Similarity 96.2%; Pred. No. 7e-103;
Matches 204; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 GRLLS-----GQPVCRGGTQPCYKVIYFHTSRLNFEEAKEACRRDGGQLVSIES 52
Db |||||
24 GRLLSASDLRLRGQPVCRGGTQPCYKVIYFHTSRLNFEEAKEACRRDGGQLVSIES 83

QY 53 EDEQKLIKFIENLLPSDGFWIGLRRREKQSNSTACQDLYAWTDGSIQFRNWWYVDEP 112
Db |||||
84 EDEQKLIKFIENLLPSDGFWIGLRRREKQSNSTACQDLYAWTDGSIQFRNWWYVDEP 143

QY 113 SCGSEVCVVMYHQPSAPAGIGGPFQWDDRCNMKNFICKYSDEKPAVPSREAEGEET 172
Db |||||
144 SCGSEVCVVMYHQPSAPAGIGGPFQWDDRCNMKNFICKYSDEKPAVPSREAEGEET 203

QY 173 ELTTPVLPEETQBEDAKKTFKESREAAALNLAY 204

Db 204 ELTTPVLPEETQBEDAKKTFKESREAAALNLAY 235
|||
RESULT 8
AAU29033
ID AAU29033 standard; Protein; 382 AA.
XX
AC AAU29033;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human PRO polypeptide sequence #10.
XX
KW PRO polypeptide; mammal; tumour; cancer; human; cattle; horse; sheep;
KW dog; cat; pig; goat; rabbit; tumour necrosis factor alpha; TNF-alpha;
KW blood; chondrocyte cell; cell proliferation; cell differentiation; colon;
KW adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200168848-A2.
XX
PD 20-SEP-2001.
XX
PF 28-FEB-2001; 2001WO-US06520.
XX
PR 01-MAR-2000; 2000WO-US05601.
PR 02-MAR-2000; 2000WO-US05841.
PR 03-MAR-2000; 2000US-187202P.
PR 06-MAR-2000; 2000US-186968P.
PR 14-MAR-2000; 2000US-189320P.
PR 14-MAR-2000; 2000US-189328P.
PR 15-MAR-2000; 2000WO-US06884.
PR 21-MAR-2000; 2000US-190828P.
PR 21-MAR-2000; 2000US-191007P.
PR 21-MAR-2000; 2000US-191048P.
PR 21-MAR-2000; 2000US-191314P.
PR 28-MAR-2000; 2000US-192655P.
PR 29-MAR-2000; 2000US-193032P.
PR 29-MAR-2000; 2000US-193053P.
PR 30-MAR-2000; 2000WO-US08439.
PR 04-APR-2000; 2000US-194449P.
PR 04-APR-2000; 2000US-194647P.
PR 11-APR-2000; 2000US-195975P.
PR 11-APR-2000; 2000US-196000P.
PR 11-APR-2000; 2000US-196187P.
PR 11-APR-2000; 2000US-196690P.
PR 11-APR-2000; 2000US-196820P.
PR 18-APR-2000; 2000US-198121P.
PR 18-APR-2000; 2000US-198585P.
PR 25-APR-2000; 2000US-199397P.
PR 25-APR-2000; 2000US-199550P.
PR 25-APR-2000; 2000US-199654P.
PR 03-MAY-2000; 2000US-201516P.
PR 17-MAY-2000; 2000WO-US13705.
PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.
PR 02-JUN-2000; 2000WO-US15264.
PR 05-JUN-2000; 2000US-209832P.
PR 28-JUL-2000; 2000WO-US20710.
PR 22-AUG-2000; 2000US-0644848.
PR 24-AUG-2000; 2000WO-US23328.
PR 08-NOV-2000; 2000WO-US30952.
PR 01-DEC-2000; 2000WO-US32678.
PR 20-DEC-2000; 2000WO-US34956.
XX
(GETH) GENENTECH INC.
PA
XX Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;
PI Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;
XX WPI; 2001-602746/68.
DR

DR N-PSDB; AAS45934.
XX Novel nucleic acids encoding PRO polypeptides, used to diagnose the
PT presence of tumours, such as prostate and breast tumours, in mammals and
PT to screen for modulators of the compounds -
XX
PS Claim 11; Fig 20; 774pp; English.
XX
CC Sequences AAU29024-AAU29328 represent PRO polypeptides of the invention.
CC The PRO polypeptides and their associated nucleic acids can be used to
CC detect the presence of a tumour in a mammal by comparing the level of
CC expression of a PRO polypeptide in a test sample of cells from the animal
CC and a control sample of normal cells, whereby a higher level of
CC expression in the test sample indicates the presence of a tumour in the
CC mammal. Mammals include dogs, cats, cattle, horses, sheep, pigs, goats
CC and rabbits but are preferably human. The polypeptides can be used to
CC stimulate tumour necrosis factor (TNF) alpha release from human blood,
CC when contacted with it. A specific polypeptide can be used to stimulate
CC the proliferation or differentiation of chondrocyte cells. The PRO
CC proteins can be used to determine the presence of tumours and also
CC susceptibility to tumour development, particularly adrenal, lung, colon,
CC breast, prostate, rectal, cervical, or liver tumours, in mammalian
CC subjects. The oligonucleotide probes specific for the PRO nucleic acids
CC can be used for genetic analysis of individuals with genetic disorders.
XX
SQ Sequence 382 AA;

Query Match 98.7%; Score 1092; DB 22; Length 382;
Best Local Similarity 96.2%; Pred. No. 7e-103;
Matches 204; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
QY 1 GRLLS-----GQVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIES 52
Db |||||
24 GRLLSASDLRLGGQPVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIES 83
QY 53 EDEQKLIKFIENLLPSDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNYYVDEP 112
Db |||||
84 EDEQKLIKFIENLLPSDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNYYVDEP 143
QY 113 SCGSEVCVMYHQPSAPAGIGGPMFQWDDRCNMKNFICKYSDEKPAVPSREAEGEET 172
Db |||||
144 SCGSEVCVMYHQPSAPAGIGGPMFQWDDRCNMKNFICKYSDEKPAVPSREAEGEET 203
QY 173 ELTTPVLPEETQEEADAKTKFKESREAAALNLAY 204
Db |||||
204 ELTTPVLPEETQEEADAKTKFKESREAAALNLAY 235

RESULT 9
AAB80235
ID AAB80235 standard; Protein; 382 AA.
XX
AC AAB80235;
DT 24-APR-2001 (first entry)
XX
DE Human PRO234 protein.
KW Human; PRO; dermatological; antipsoriatic; cytostatic; antiinflammatory;
KW antiparkinsonian nootropic; neuroprotective; vulnerary; cardiant;
KW antiangiogenic; vasotropic; antiasthmatic; antirheumatic; cancer;
KW antiarthritic; antiinfertility; antidiabetic; antiviral; diabetes;
KW ophthalmological; gene therapy; skin disease; gastrointestinal disorder;
KW ischaemia; inflammation.
XX
OS Homo sapiens.
XX
PN WO200104311-A1.
XX
PD 18-JAN-2001.
XX
PF 22-FEB-2000; 2000WO-US04414.
XX

PR 07-JUL-1999; 99US-0143048.
PR 26-JUL-1999; 99US-0145698.
PR 28-JUL-1999; 99US-0146222.
PR 08-SEP-1999; 99WO-US20594.
PR 13-SEP-1999; 99WO-US20944.
PR 15-SEP-1999; 99WO-US21090.
PR 15-SEP-1999; 99WO-US21547.
PR 05-OCT-1999; 99WO-US23089.
PR 29-NOV-1999; 99WO-US28214.
PR 30-NOV-1999; 99WO-US28313.
PR 16-DEC-1999; 99WO-US30095.
PR 20-DEC-1999; 99WO-US30911.
PR 20-DEC-1999; 99WO-US30999.
PR 05-JAN-2000; 99WO-US00219.
XX
XX (GETH) GENENTECH INC.
PA Ashkenazi AJ, Botstein D, Desnoyers L, Eaton DL, Ferrara N;
PI Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME, Goddard A;
PI Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ, Kljavin IJ;
PI Mather JP, Pan J, Paoni NF, Roy MA, Stewart TA, Tumas D;
PI Williams PM, Wood WI;
XX
DR WPI; 2001-081051/09.
DR N-PSDB; AAF72396.
XX
PT Sixty one nucleic acids encoding PRO polypeptides which are useful in
PT the treatment of skin diseases (e.g. psoriasis), cancers (e.g. lung
PT squamous cell carcinoma) and neurodegenerative diseases (e.g.
PT Alzheimer's disease) -
XX
PS Claim 1; Fig 50; 393pp; English.
XX
CC The present sequence is one of sixty one novel secreted and
CC transmembrane PRO polypeptides. The PRO polypeptides are
CC useful for treating skin diseases (e.g. psoriasis), cancers (e.g. lung
CC squamous cell carcinoma), gastrointestinal disorders (e.g.
CC enterocolitis), neurodegenerative diseases (e.g. Alzheimer's disease,
CC Parkinson's disease), wound repair, cardiovascular disorders (e.g.
CC endometrial bleeding angiogenesis, ischaemias such as coronary
CC ischaemia, atherosclerosis), inflammatory disorders (e.g. asthma,
CC rheumatoid arthritis, multiple sclerosis), infertility, AIDS and
CC diabetes and retinal disorders such as retinitis pigmentosum.
CC The PRO nucleic acids have applications in molecular biology, including
CC use as hybridization probes, and in chromosome and gene mapping.
XX
SQ Sequence 382 AA;

Query Match 98.7%; Score 1092; DB 22; Length 382;
Best Local Similarity 96.2%; Pred. No. 7e-103;
Matches 204; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
QY 1 GRLLS-----GQVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIES 52
Db |||||
24 GRLLSASDLRLGGQPVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIES 83
QY 53 EDEQKLIKFIENLLPSDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNYYVDEP 112
Db |||||
84 EDEQKLIKFIENLLPSDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNYYVDEP 143
QY 113 SCGSEVCVMYHQPSAPAGIGGPMFQWDDRCNMKNFICKYSDEKPAVPSREAEGEET 172
Db |||||
144 SCGSEVCVMYHQPSAPAGIGGPMFQWDDRCNMKNFICKYSDEKPAVPSREAEGEET 203
QY 173 ELTTPVLPEETQEEADAKTKFKESREAAALNLAY 204
Db |||||
204 ELTTPVLPEETQEEADAKTKFKESREAAALNLAY 235

RESULT 10
ABU69645
ID ABU69645 standard; Protein; 382 AA.
XX

AC ABU69645;
XX
DT 05-JUN-2003 (first entry)
XX
DE Novel human secreted and transmembrane protein PRO234.
XX
KW Human; secreted and transmembrane protein; gene therapy; psoriasis;
KW enterocolitis; gastrointestinal ulceration; skin disease;
KW keratinocyte differentiation; epithelial cancer; Alzheimer's disease;
KW squamous cell carcinoma; Parkinson's disease; inflammatory disease;
KW amyotrophic lateral sclerosis; rheumatoid arthritis; asthma;
KW multiple sclerosis; organ failure; atherosclerosis; cardiac injury;
KW infertility; birth defect; premature aging; AIDS; cancer;
XX diabetic complication; wound repair; tissue re-growth.
OS Homo sapiens.
XX
XX US2003017463-A1.
PN
XX
XX 23-JAN-2003.
PD
XX
XX 11-JUL-2001; 2001US-0903640.
PF
XX 10-SEP-1998; 98WO-US18824.
XX 14-SEP-1998; 98WO-US19177.
PR 16-SEP-1998; 98WO-US19330.
PR 17-SEP-1998; 98WO-US19437.
PR 01-DEC-1998; 98WO-US25108.
PR 08-SEP-1999; 99WO-US20594.
PR 13-SEP-1999; 99WO-US20944.
PR 15-SEP-1999; 99WO-US21090.
PR 15-SEP-1999; 99WO-US21547.
PR 05-OCT-1999; 99WO-US23089.
PR 29-NOV-1999; 99WO-US28214.
PR 30-NOV-1999; 99WO-US28313.
PR 01-DEC-1999; 99WO-US28301.
PR 02-DEC-1999; 99WO-US28564.
PR 02-DEC-1999; 99WO-US28565.
PR 16-DEC-1999; 99WO-US30095.
PR 20-DEC-1999; 99WO-US30911.
PR 20-DEC-1999; 99WO-US30999.
PR 05-JAN-2000; 2000WO-US00219.
PR 11-FEB-2000; 2000WO-US03565.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US05004.
PR 02-MAR-2000; 2000WO-US05841.
PR 20-MAR-2000; 2000WO-US07377.
PR 30-MAR-2000; 2000WO-US08439.
PR 22-MAY-2000; 2000WO-US14042.
PR 02-JUN-2000; 2000WO-US15264.
PR 28-JUL-2000; 2000WO-US20710.
PR 24-AUG-2000; 2000WO-US23328.
PR 17-SEP-1997; 97US-059113P.
PR 17-SEP-1997; 97US-059115P.
PR 17-SEP-1997; 97US-059117P.
PR 17-SEP-1997; 97US-059119P.
PR 17-SEP-1997; 97US-059121P.
PR 17-SEP-1997; 97US-059122P.
PR 17-SEP-1997; 97US-059184P.
PR 18-SEP-1997; 97US-059263P.
PR 18-SEP-1997; 97US-059266P.
PR 15-OCT-1997; 97US-062125P.
PR 17-OCT-1997; 97US-062285P.
PR 17-OCT-1997; 97US-062287P.
PR 21-OCT-1997; 97US-063486P.
PR 24-OCT-1997; 97US-062814P.
PR 24-OCT-1997; 97US-062816P.
PR 24-OCT-1997; 97US-063045P.
PR 24-OCT-1997; 97US-063120P.
PR 24-OCT-1997; 97US-063121P.
PR 24-OCT-1997; 97US-063127P.
PR 24-OCT-1997; 97US-063128P.
PR 27-OCT-1997; 97US-063327P.

PR 27-OCT-1997; 97US-063329P.
PR 28-OCT-1997; 97US-063541P.
PR 28-OCT-1997; 97US-063542P.
PR 28-OCT-1997; 97US-063544P.
PR 28-OCT-1997; 97US-063549P.
PR 28-OCT-1997; 97US-063550P.
PR 28-OCT-1997; 97US-063564P.
PR 29-OCT-1997; 97US-063435P.
PR 29-OCT-1997; 97US-063704P.
PR 29-OCT-1997; 97US-063732P.
PR 29-OCT-1997; 97US-063734P.
PR 29-OCT-1997; 97US-063735P.
PR 29-OCT-1997; 97US-063738P.
PR 29-OCT-1997; 97US-064215P.
PR 31-OCT-1997; 97US-063870P.
PR 31-OCT-1997; 97US-064103P.
PR 03-NOV-1997; 97US-064248P.
PR 07-NOV-1997; 97US-064809P.
PR 12-NOV-1997; 97US-065186P.
PR 17-NOV-1997; 97US-065846P.
PR 18-NOV-1997; 97US-065693P.
PR 21-NOV-1997; 97US-066120P.
PR 21-NOV-1997; 97US-066364P.
PR 24-NOV-1997; 97US-066453P.
PR 24-NOV-1997; 97US-066466P.
PR 24-NOV-1997; 97US-066511P.
PR 24-NOV-1997; 97US-066770P.
PR 24-NOV-1997; 97US-066772P.
PR 25-NOV-1997; 97US-066840P.
PR 12-DEC-1997; 97US-069425P.
PR 04-JUN-1998; 98US-088026P.
PR 10-SEP-1998; 98US-099803P.
PR 14-SEP-1998; 98US-100262P.
PR 17-SEP-1998; 98US-100858P.
PR 13-OCT-1998; 98US-104080P.
PR 20-NOV-1998; 98US-109304P.
PR 22-DEC-1998; 98US-113296P.
PR 07-JUL-1999; 99US-143048P.
PR 26-JUL-1999; 99US-145698P.
PR 28-JUL-1999; 99US-146222P.
PR 18-SEP-2000; 2000US-0665350.
XX
PA (GETH) GENENTECH INC.

Ashkenazi A, Botstein D, Desnoyers L, Eaton DL, Ferrara N;
Filvaroff E, Fong S, Gerber H, Gerritsen ME, Goddard A;
Godowski PJ, Grimaldi JC, Gurney AL, Hillan KJ, Kljavin IJ;
Mather JP, Pan J, Paoni NF, Roy MA, Stewart TA, Tumas D;
Williams PM, Wood WI;

WPI: 2003-341586/32.
N-PSDB; ACA54920.

New PRO polypeptides and nucleic acid molecules, useful in diagnosing
or treating inflammatory diseases, organ failure, atherosclerosis,
cardiac injury, infertility, cancer, AIDS, Alzheimer's disease or
Parkinson's disease -

Claim 12; Fig 50; 473pp; English.

The invention describes sixty one nucleic acids encoding PRO polypeptides
(secreted and transmembrane). The PRO polypeptides and nucleic acids are
useful in diagnosing or treating enterocolitis, gastrointestinal
ulceration, skin diseases associated with abnormal keratinocyte
differentiation, e.g. psoriasis or epithelial cancers such as squamous
cell carcinoma, Alzheimer's disease, Parkinson's disease, amyotrophic
lateral sclerosis, inflammatory diseases, e.g. rheumatoid arthritis,
asthma or multiple sclerosis, organ failure, atherosclerosis, cardiac
injury, infertility, birth defects, premature aging, AIDS, cancer,
diabetic complications, or mutations in general. The polypeptides are
also useful for wound repair and associated therapies concerned with
re-growth of tissue. The PRO polypeptides and nucleic acid molecules
are also useful in gene therapy, and as molecular weight markers for

CC protein electrophoresis purposes. The anti-PRO antibodies may be used
CC in diagnostic assays for PRO, or for the affinity purification of PRO
CC from recombinant cell culture or natural sources. This is the amino
CC acid sequence of a novel human PRO polypeptide.
XX
SQ Sequence 382 AA;

Query Match 98.7%; Score 1092; DB 24; Length 382;
Best Local Similarity 96.2%; Pred. No. 7e-103;
Matches 204; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 GRLLS-----GQVCRGGTQPCYKVIYFHDTSRRLNFEAKEACRRDGGQLVSIES 52
Dd 24 GRLLSASDLDRGGQPVCRGGTQPCYKVIYFHDTSRRLNFEAKEACRRDGGQLVSIES 83

QY 53 EDEQKLIKFIENLLPSDGFWIGLRRRREKQSNSTACQDLYAWTDGSIQFRNWWYDEP 112
Dd 84 EDEQKLIKFIENLLPSDGFWIGLRRRREKQSNSTACQDLYAWTDGSIQFRNWWYDEP 143

QY 113 SCGSEVCVMYHQPAPAGIGGPFQWDDRCNMKNFICKYSDEKPAVPSREAGEET 172
Dd 144 SCGSEVCVMYHQPAPAGIGGPFQWDDRCNMKNFICKYSDEKPAVPSREAGEET 203

QY 173 ELTPVLPETQEEADAKTFKESREAAALNLAY 204
Dd 204 ELTPVLPETQEEADAKTFKESREAAALNLAY 235

RESULT 11
ABU71121
ID ABU71121 standard; Protein; 382 AA.
XX
AC ABU71121;
XX
DT 10-JUN-2003 (first entry)
XX
DE Human PRO234 protein.
XX
KW Human; PRO; secreted; transmembrane; cytostatic; TNF-alpha; blood;
KW tumour necrosis factor alpha release; chondrocyte cell; proliferation;
KW differentiation; tumour; gene therapy.
XX
OS Homo sapiens.
XX
PN UF2003036143-A1.
XX
PD 20-FEB-2003.
XX
PF 02-JUL-2002; 2002US-0187600.
XX
PR 16-SEP-1998; 98WO-US19330.
PR 07-OCT-1998; 98WO-US21141.
PR 01-DEC-1998; 98WO-US25108.
PR 08-MAR-1999; 99WO-US05028.
PR 14-MAY-1999; 99WO-US10733.
PR 02-JUN-1999; 99WO-US12252.
PR 01-SEP-1999; 99WO-US20111.
PR 15-SEP-1999; 99WO-US21090.
PR 01-DEC-1999; 99WO-US28301.
PR 02-DEC-1999; 99WO-US28551.
PR 30-DEC-1999; 99WO-US31274.
PR 05-JAN-2000; 2000WO-US00219.
PR 18-FEB-2000; 2000WO-US04341.
PR 18-FEB-2000; 2000WO-US04342.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US05004.
PR 01-MAR-2000; 2000WO-US05601.
PR 02-MAR-2000; 2000WO-US05841.
PR 15-MAR-2000; 2000WO-US06884.
PR 30-MAR-2000; 2000WO-US08439.
PR 17-MAY-2000; 2000WO-US13705.
PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.

PR 02-JUN-2000; 2000WO-US15264.
PR 28-JUL-2000; 2000WO-US20710.
PR 24-AUG-2000; 2000WO-US23328.
PR 08-NOV-2000; 2000WO-US30952.
PR 01-DEC-2000; 2000WO-US32678.
PR 20-DEC-2000; 2000WO-US34956.
PR 28-FEB-2001; 2001WO-US06520.
PR 01-JUN-2001; 2001WO-US17800.
PR 20-JUN-2001; 2001WO-US19692.
PR 29-JUN-2001; 2001WO-US21066.
PR 09-JUL-2001; 2001WO-US21735.
PR 29-AUG-2001; 2001WO-US27099.
PR 18-SEP-1997; 97US-059263P.
PR 18-SEP-1997; 97US-059266P.
PR 17-OCT-1997; 97US-062250P.
PR 21-OCT-1997; 97US-063486P.
PR 24-OCT-1997; 97US-063120P.
PR 24-OCT-1997; 97US-063121P.
PR 28-OCT-1997; 97US-063540P.
PR 28-OCT-1997; 97US-063541P.
PR 28-OCT-1997; 97US-063544P.
PR 28-OCT-1997; 97US-063564P.
PR 29-OCT-1997; 97US-063734P.
PR 31-OCT-1997; 97US-063870P.
PR 31-OCT-1997; 97US-064103P.
PR 13-NOV-1997; 97US-065311P.
PR 21-NOV-1997; 97US-066120P.
PR 24-NOV-1997; 97US-066466P.
PR 24-NOV-1997; 97US-066772P.
PR 11-DEC-1997; 97US-069335P.
PR 12-DEC-1997; 97US-069425P.
PR 17-DEC-1997; 97US-069870P.
PR 18-DEC-1997; 97US-068017P.
PR 10-MAR-1998; 98US-077450P.
PR 11-MAR-1998; 98US-077632P.
PR 20-MAR-1998; 98US-077649P.
PR 20-MAR-1998; 98US-078886P.
PR 20-MAR-1998; 98US-078939P.
PR 27-MAR-1998; 98US-079664P.
PR 27-MAR-1998; 98US-079786P.
PR 31-MAR-1998; 98US-080107P.
PR 31-MAR-1998; 98US-080194P.
PR 01-APR-1998; 98US-080327P.
PR 01-APR-1998; 98US-080333P.
PR 08-APR-1998; 98US-081049P.
PR 08-APR-1998; 98US-081070P.
PR 09-APR-1998; 98US-081195P.
PR 15-APR-1998; 98US-081838P.
PR 21-APR-1998; 98US-082568P.
PR 21-APR-1998; 98US-082569P.
PR 22-APR-1998; 98US-082704P.
PR 22-APR-1998; 98US-082797P.
PR 28-APR-1998; 98US-083322P.
PR 29-APR-1998; 98US-083495P.
PR 29-APR-1998; 98US-083496P.
PR 29-APR-1998; 98US-083499P.
PR 29-APR-1998; 98US-083559P.
PR 05-MAY-1998; 98US-084366P.
PR 06-MAY-1998; 98US-084414P.
PR 07-MAY-1998; 98US-084639P.
PR 07-MAY-1998; 98US-084640P.
PR 07-MAY-1998; 98US-084643P.
PR 15-MAY-1998; 98US-085579P.
PR 15-MAY-1998; 98US-085580P.
PR 15-MAY-1998; 98US-085582P.
PR 15-MAY-1998; 98US-085700P.
PR 18-MAY-1998; 98US-086023P.
PR 22-MAY-1998; 98US-086392P.
PR 22-MAY-1998; 98US-086486P.
PR 28-MAY-1998; 98US-087098P.
PR 28-MAY-1998; 98US-087208P.
PR 02-JUN-1998; 98US-087609P.
PR 02-JUN-1998; 98US-087759P.

PR 03-JUN-1998; 98US-087827P.
PR 04-JUN-1998; 98US-088025P.
PR 04-JUN-1998; 98US-088028P.
PR 04-JUN-1998; 98US-088029P.
PR 04-JUN-1998; 98US-088033P.
PR 04-JUN-1998; 98US-088326P.
PR 05-JUN-1998; 98US-088167P.
PR 05-JUN-1998; 98US-088202P.
PR 05-JUN-1998; 98US-088212P.
PR 05-JUN-1998; 98US-088217P.
PR 09-JUN-1998; 98US-088655P.
PR 10-JUN-1998; 98US-088722P.
PR 10-JUN-1998; 98US-088738P.
PR 10-JUN-1998; 98US-088740P.
PR 10-JUN-1998; 98US-088811P.
PR 10-JUN-1998; 98US-088824P.
PR 10-JUN-1998; 98US-088825P.
PR 10-JUN-1998; 98US-088826P.
PR 11-JUN-1998; 98US-088861P.
PR 11-JUN-1998; 98US-088863P.
PR 11-JUN-1998; 98US-088876P.
PR 12-JUN-1998; 98US-089090P.
PR 12-JUN-1998; 98US-089105P.
PR 16-JUN-1998; 98US-089512P.
PR 16-JUN-1998; 98US-089514P.
PR 17-JUN-1998; 98US-089538P.
PR 17-JUN-1998; 98US-089598P.
PR 17-JUN-1998; 98US-089653P.
PR 18-JUN-1998; 98US-089908P.
PR 19-JUN-1998; 98US-089952P.
PR 22-JUN-1998; 98US-090246P.
PR 22-JUN-1998; 98US-090252P.
PR 22-JUN-1998; 98US-090254P.
PR 24-JUN-1998; 98US-090429P.
PR 24-JUN-1998; 98US-090435P.
PR 24-JUN-1998; 98US-090444P.
PR 24-JUN-1998; 98US-090461P.
PR 24-JUN-1998; 98US-090535P.
PR 24-JUN-1998; 98US-090540P.
PR 25-JUN-1998; 98US-090676P.
PR 25-JUN-1998; 98US-090678P.
PR 25-JUN-1998; 98US-090688P.
PR 25-JUN-1998; 98US-090690P.
PR 25-JUN-1998; 98US-090694P.
PR 25-JUN-1998; 98US-090695P.
PR 25-JUN-1998; 98US-090696P.
PR 26-JUN-1998; 98US-090862P.
PR 26-JUN-1998; 98US-090863P.
PR 26-JUN-1998; 98US-091010P.
PR 01-JUL-1998; 98US-091359P.
PR 01-JUL-1998; 98US-091544P.
PR 02-JUL-1998; 98US-091478P.
PR 02-JUL-1998; 98US-091486P.
PR 02-JUL-1998; 98US-091626P.
PR 02-JUL-1998; 98US-091628P.
PR 02-JUL-1998; 98US-091632P.
PR 24-JUL-1998; 98US-094006P.
PR 04-AUG-1998; 98US-095282P.
PR 10-AUG-1998; 98US-095998P.
PR 10-AUG-1998; 98US-096012P.
PR 17-AUG-1998; 98US-096757P.
PR 17-AUG-1998; 98US-096766P.
PR 17-AUG-1998; 98US-096867P.
PR 17-AUG-1998; 98US-096891P.
PR 17-AUG-1998; 98US-096897P.
PR 18-AUG-1998; 98US-096949P.
PR 18-AUG-1998; 98US-096959P.
PR 18-AUG-1998; 98US-097022P.
PR 26-AUG-1998; 98US-097952P.
PR 26-AUG-1998; 98US-097954P.
PR 26-AUG-1998; 98US-097955P.
PR 26-AUG-1998; 98US-097971P.
PR 26-AUG-1998; 98US-097974P.

PR 26-AUG-1998; 98US-098014P.
PR 01-SEP-1998; 98US-098716P.
PR 01-SEP-1998; 98US-098723P.
PR 02-SEP-1998; 98US-098803P.
PR 02-SEP-1998; 98US-098821P.
PR 02-SEP-1998; 98US-098843P.
PR 09-SEP-1998; 98US-099602P.
PR 10-SEP-1998; 98US-099741P.
PR 10-SEP-1998; 98US-099754P.
PR 10-SEP-1998; 98US-099763P.
PR 10-SEP-1998; 98US-099812P.

Query Match 98.7%; Score 1092; DB 24; Length 382;
Best Local Similarity 96.2%; Pred. No. 7e-103;
Matches 204; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

Qy 1 GRLLS-----GQVCRGGTQPCYKVIYFHDTSRRLNFEFEAKEACRRDGGQLVSIES 52
Db 24 GRLLSASDLRLGGQPVCRGGTQPCYKVIYFHDTSRRLNFEFEAKEACRRDGGQLVSIES 83

Qy 53 EDEQKLIKFIEIENLLPSDGFWIGLRRREBKQSNSTACQDLYAWTDGSI SQFRNMYVDEP 112
Db 84 EDEQKLIKFIEIENLLPSDGFWIGLRRREBKQSNSTACQDLYAWTDGSI SQFRNMYVDEP 143

Qy 113 SCGSEVCVMYHQPSAPAGIGGPFQWDDRCNMKNPFICKYSDEKPAVPSREAEGET 172
Db 144 SCGSEVCVMYHQPSAPAGIGGPFQWDDRCNMKNPFICKYSDEKPAVPSREAEGET 203

Qy 173 ELTPVLPETQEEDAKKTFKESREAAALNLAY 204
Db 204 ELTPVLPETQEEDAKKTFKESREAAALNLAY 235

RESULT 12

ABU71468
ID ABU71468 standard; Protein; 382 AA.
XX
AC ABU71468;
XX
DT 10-JUN-2003 (first entry)
XX
DE Human PRO polypeptide #24.
XX
KW Human; secreted and transmembrane protein; PRO polypeptide; cancer;
KW Alzheimer's disease; ischaemia; cytostatic; nootropic; vasotropic;
KW neuroprotective.
XX
OS Homo sapiens.
XX
PN US2002192659-A1.
XX
PD 19-DEC-2002.
XX
PF 10-JUL-2001; 2001US-0902853.
XX
PR 10-SEP-1998; 98WO-US18824.
PR 14-SEP-1998; 98WO-US19177.
PR 16-SEP-1998; 98WO-US19330.
PR 17-SEP-1998; 98WO-US19437.
PR 01-DEC-1998; 98WO-US25108.
PR 08-SEP-1999; 99WO-US20594.
PR 13-SEP-1999; 99WO-US20944.
PR 15-SEP-1999; 99WO-US21090.
PR 15-SEP-1999; 99WO-US21547.
PR 05-OCT-1999; 99WO-US23089.
PR 01-DEC-1999; 99WO-US28301.
PR 02-DEC-1999; 99WO-US28564.
PR 02-DEC-1999; 99WO-US28565.
PR 16-DEC-1999; 99WO-US30095.
PR 20-DEC-1999; 99WO-US30911.
PR 20-DEC-1999; 99WO-US30999.
PR 05-JAN-2000; 2000WO-US00219.
PR 11-FEB-2000; 2000WO-US03565.

PR 22-FEB-2000; 200WO-US04414.
PR 28-JUL-2000; 200WO-US20710.
PR 24-AUG-2000; 200WO-US23328.
PR 17-SEP-1997; 97US-059113P.
PR 17-SEP-1997; 97US-059115P.
PR 17-SEP-1997; 97US-059117P.
PR 18-SEP-1997; 97US-059266P.
PR 15-OCT-1997; 97US-062125P.
PR 17-OCT-1997; 97US-062285P.
PR 17-OCT-1997; 97US-062287P.
PR 21-OCT-1997; 97US-063486P.
PR 24-OCT-1997; 97US-062814P.
PR 24-OCT-1997; 97US-062816P.
XX

PA (GETH) GENENTECH INC.

XX
PI Ashkenazi A, Botstein D, Desnoyers L, Eaton DL, Ferrara N;
PI Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME, Goddard A;
PI Godowski PJ, Grimaldi JC, Gurney AL, Hillan KJ, Kljavin IJ;
PI Mather JP, Pan J, Paoni NF, Roy MA, Stewart TA, Tumas D;
PI Williams PM, Wood WI;
XX

DR WPI; 2003-361832/34.
DR N-PSDB; ACA58405.

XX
PT New isolated nucleic acid encoding a PRO polypeptide, e.g. PRO245 or
PT PRO1868, useful in molecular biology, chromosome and gene mapping, in
PT generating antisense RNA and DNA, and in gene therapy -
XX

PS Claim 12; Fig 50; 474pp; English.

XX
CC The present invention relates to the isolation of novel human secreted
CC and transmembrane proteins (PRO polypeptides), and the polynucleotide
CC sequences encoding them. The polynucleotide sequences are useful in
CC molecular biology, as hybridisation probes, in chromosome and gene
CC mapping, in generating antisense RNA and DNA, and in gene therapy. The
CC polynucleotide sequences may also be used in preparing PRO polypeptides
CC by recombinant techniques, and in generating either transgenic animals
CC or knock-out animals which, in turn, are useful in the development and
CC screening of therapeutically useful reagents. The PRO polypeptides or
CC their antibodies are useful in preparing a medicament for treating a
CC condition responsive to the polypeptide or antibody, such as cancer,
CC Alzheimer's disease or ischaemia, and in various diagnostic assays.
CC ABU71445-ABU71505 represent human PRO polypeptides of the invention.
XX

SQ Sequence 382 AA;

Query Match 98.7%; Score 1092; DB 24; Length 382;
Best Local Similarity 96.2%; Pred. No. 7e-103;
Matches 204; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 GRLLS-----GQVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIES 52
Db |||||
24 GRLLSASDLDRGGQPVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIES 83
QY 53 EDEQKLIKFIENLLPSDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNWWYVDEP 112
Db |||||
84 EDEQKLIKFIENLLPSDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNWWYVDEP 143
QY 113 SCGSEVCVMYHQPSAPAGIGGPPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGEET 172
Db |||||
144 SCGSEVCVMYHQPSAPAGIGGPPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGEET 203
QY 173 ELTTPVLPEETQEEDAKKTFKESREAAALNLAY 204
Db |||||
204 ELTTPVLPEETQEEDAKKTFKESREAAALNLAY 235

RESULT 13

ABU71914

ID ABU71914 standard; Protein; 382 AA.

XX

AC ABU71914;

XX
DT 12-JUN-2003 (first entry)
XX
DE Human secreted/transmembrane protein PRO234.
XX
KW Human; secreted protein; transmembrane protein; PRO;
KW gene therapy; chromosome identification; chromosome marker.
XX
OS Homo sapiens.
XX
PN US2003003530-A1.
XX
PD 02-JAN-2003.
XX
PF 11-JUL-2001; 2001US-0904011.
XX
PR 10-SEP-1998; 98WO-US18824.
PR 14-SEP-1998; 98WO-US19177.
PR 16-SEP-1998; 98WO-US19330.
PR 17-SEP-1998; 98WO-US19437.
PR 01-DEC-1998; 98WO-US25108.
PR 08-SEP-1999; 99WO-US20594.
PR 13-SEP-1999; 99WO-US20944.
PR 15-SEP-1999; 99WO-US21090.
PR 15-SEP-1999; 99WO-US21547.
PR 05-OCT-1999; 99WO-US23089.
PR 29-NOV-1999; 99WO-US28214.
PR 30-NOV-1999; 99WO-US28313.
PR 01-DEC-1999; 99WO-US28301.
PR 02-DEC-1999; 99WO-US28564.
PR 02-DEC-1999; 99WO-US28565.
PR 16-DEC-1999; 99WO-US30095.
PR 20-DEC-1999; 99WO-US30911.
PR 20-DEC-1999; 99WO-US30999.
PR 05-JAN-2000; 2000WO-US00219.
PR 11-FEB-2000; 2000WO-US03565.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US05004.
PR 02-MAR-2000; 2000WO-US05841.
PR 20-MAR-2000; 2000WO-US07377.
PR 30-MAR-2000; 2000WO-US08439.
PR 22-MAY-2000; 2000WO-US14042.
PR 02-JUN-2000; 2000WO-US15264.
PR 28-JUL-2000; 2000WO-US20710.
PR 24-AUG-2000; 2000WO-US23328.
PR 17-SEP-1997; 97US-059113P.
PR 17-SEP-1997; 97US-059115P.
PR 17-SEP-1997; 97US-059117P.
PR 17-SEP-1997; 97US-059119P.
PR 17-SEP-1997; 97US-059121P.
PR 17-SEP-1997; 97US-059122P.
PR 17-SEP-1997; 97US-059184P.
PR 18-SEP-1997; 97US-059263P.
PR 18-SEP-1997; 97US-059266P.
PR 15-OCT-1997; 97US-062125P.
PR 17-OCT-1997; 97US-062285P.
PR 17-OCT-1997; 97US-062287P.
PR 21-OCT-1997; 97US-063486P.
PR 24-OCT-1997; 97US-062814P.
PR 24-OCT-1997; 97US-062816P.
PR 24-OCT-1997; 97US-063045P.
PR 24-OCT-1997; 97US-063120P.
PR 24-OCT-1997; 97US-063121P.
PR 24-OCT-1997; 97US-063127P.
PR 24-OCT-1997; 97US-063128P.
PR 27-OCT-1997; 97US-063327P.
PR 27-OCT-1997; 97US-063329P.
PR 28-OCT-1997; 97US-063541P.
PR 28-OCT-1997; 97US-063542P.
PR 28-OCT-1997; 97US-063544P.
PR 28-OCT-1997; 97US-063549P.
PR 28-OCT-1997; 97US-063550P.
PR 28-OCT-1997; 97US-063564P.

PR 29-OCT-1997; 97US-063435P.
PR 29-OCT-1997; 97US-063704P.
PR 29-OCT-1997; 97US-063732P.
PR 29-OCT-1997; 97US-063734P.
PR 29-OCT-1997; 97US-063735P.
PR 29-OCT-1997; 97US-063738P.
PR 29-OCT-1997; 97US-064215P.
PR 31-OCT-1997; 97US-063870P.
PR 31-OCT-1997; 97US-064103P.
PR 03-NOV-1997; 97US-064248P.
PR 07-NOV-1997; 97US-064809P.
PR 12-NOV-1997; 97US-065186P.
PR 17-NOV-1997; 97US-065846P.
PR 18-NOV-1997; 97US-065693P.
PR 21-NOV-1997; 97US-066120P.
PR 21-NOV-1997; 97US-066364P.
PR 24-NOV-1997; 97US-066453P.
PR 24-NOV-1997; 97US-066466P.
PR 24-NOV-1997; 97US-066511P.
PR 24-NOV-1997; 97US-066770P.
PR 24-NOV-1997; 97US-066772P.
PR 18-SEP-2000; 2000US-0665350.
XX
PA (GETH) GENENTECH INC.
XX
PI Ashkenazi A, Botstein D, Desnoyers L, Eaton DL, Ferrara N;
PI Filvaroff E, Fong S, Gerber H, Gerritsen ME, Goddard A;
PI Godowski PJ, Grimaldi JC, Gurney AL, Hillan KJ, Kljavin IJ;
PI Mather JP, Pan J, Paoni NF, Roy MA, Stewart TA, Tumas D;
PI Williams PM, Wood WI;
XX
DR WPI; 2003-329602/31.
DR N-PSDB; ACA60112.
XX
PT New transmembrane polypeptides and nucleic acids encoding the
PT polypeptides, useful in gene therapy, in chromosome identification, as
PT chromosome markers, in generating probes and in tissue typing -
XX
PS Claim 12; Fig 50; 484pp; English.
XX
CC The invention relates to an isolated nucleic acid with at least 80%
CC nucleic acid sequence identity to a nucleotide sequence encoding one of
CC 61 secreted/transmembrane polypeptides, or PRO polypeptides or encoding a
CC PRO protein extracellular domain. Also included are a vector comprising
CC the PRO nucleic acid, a host cell comprising the vector, producing a PRO
CC polypeptide (by culturing the host cell for the expression of the PRO
CC polypeptide, and recovering the PRO polypeptide from the cell culture),
CC an isolated PRO polypeptide (having at least 80% sequence identity
CC to: (a) an amino acid sequence selected from the 61 PRO proteins;
CC (b) an amino acid sequence encoded by a nucleic acid molecule deposited
CC with an ATCC number (detailed in the specification); or (c) an
CC extracellular domain of a PRO polypeptide or to a PRO polypeptide lacking
CC its associated signal peptide), a chimaeric molecule comprising a PRO
CC polypeptide of fused to a heterologous amino acid sequence, an anti-PRO
CC antibody, detecting a PRO245 or PRO1868 in a sample suspected of
CC containing the polypeptide, linking a bioactive molecule to a cell
CC expressing a PRO245 or PRO1868 and modulating at least one biological
CC activity of a cell expressing a PRO245 or PRO1868. Nucleic acids which
CC encode PRO can be used to generate either transgenic animals or knock-out
CC animals which may be used in the development and screening of
CC therapeutically useful reagents. The nucleic acids may also be used in
CC gene therapy, in chromosome identification, as chromosome markers, or in
CC generating probes. The PRO polypeptides are useful as molecular markers
CC for protein electrophoresis, and the isolated nucleic acids may be used
CC for recombinantly expressing those markers. The PRO polypeptides and
CC nucleic acids may also be used in tissue typing. Anti-PRO antibodies
CC are useful in diagnostic assays for PRO, and in affinity purification
CC of PRO from recombinant cell culture or natural sources. The
XX present sequence represents a PRO protein.
SQ Sequence 382 AA;

Query Match

98.7%; Score 1092; DB 24; Length 382;

Best Local Similarity 96.2%; Pred. No. 7e-103;
Matches 204; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
QY 1 GRLLS-----GQVCRGGTQRPCYKVIYFHDTSRRLNFEAEKACRRDGGQLVSIES 52
| | | | |
DB 24 GRLLSASDLRGGQPVCRCGGTQRPCYKVIYFHDTSRRLNFEAEKACRRDGGQLVSIES 83
| | | | |
QY 53 EDEQKLIKFIENLLPSDGFWIGLRRREKQSNSTACQDLYAWTDGSIQFRNMYVDEP 112
| | | | |
DB 84 EDEQKLIKFIENLLPSDGFWIGLRRREKQSNSTACQDLYAWTDGSIQFRNMYVDEP 143
| | | | |
QY 113 SCGSEVCVMYHQPSAPAGIGGPFQWMDRCNMKNFICKYSDEKPAVPSREAEGET 172
| | | | |
DB 144 SCGSEVCVMYHQPSAPAGIGGPFQWMDRCNMKNFICKYSDEKPAVPSREAEGET 203
| | | | |
QY 173 ELTTPVLPEETQEEDAKTKFKESREAAALNLAY 204
| | | | |
DB 204 ELTTPVLPEETQEEDAKTKFKESREAAALNLAY 235
| | | | |
RESULT 14
ABU65578
ID ABU65578 standard; Protein; 382 AA.
XX
AC ABU65578;
XX
XX 19-MAY-2003 (first entry)
XX
DE Human secreted/transmembrane protein, SEQ ID 20.
XX
KW Human; PRO; secreted protein; transmembrane protein;
KW cytosstatic; antiarthritic; osteopathic; adrenal tumour; lung tumour;
KW colon tumour; breast tumour; prostate tumour; rectal tumour;
KW cervical tumour; liver tumour; TNF-alpha release; arthritis;
KW tumour necrosis factor alpha; chondrocyte cell; bone disorder;
KW cartilage disorder; sports injury.
XX
OS Homo sapiens.
XX
PN US2003036156-A1.
XX
XX 20-FEB-2003.
XX
PF 02-JUL-2002; 2002US-0188767.
XX
PR 16-SEP-1998; 98WO-US19330.
PR 07-OCT-1998; 98WO-US21141.
PR 01-DEC-1998; 98WO-US25108.
PR 08-MAR-1999; 99WO-US05028.
PR 14-MAY-1999; 99WO-US10733.
PR 02-JUN-1999; 99WO-US12252.
PR 01-SEP-1999; 99WO-US20111.
PR 15-SEP-1999; 99WO-US21090.
PR 01-DEC-1999; 99WO-US28301.
PR 02-DEC-1999; 99WO-US28551.
PR 30-DEC-1999; 99WO-US31274.
PR 05-JAN-2000; 2000WO-US00219.
PR 18-FEB-2000; 2000WO-US04341.
PR 18-FEB-2000; 2000WO-US04342.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US05004.
PR 01-MAR-2000; 2000WO-US05601.
PR 02-MAR-2000; 2000WO-US05841.
PR 15-MAR-2000; 2000WO-US06884.
PR 30-MAR-2000; 2000WO-US08439.
PR 17-MAY-2000; 2000WO-US13705.
PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.
PR 02-JUN-2000; 2000WO-US15264.
PR 28-JUL-2000; 2000WO-US20710.
PR 24-AUG-2000; 2000WO-US23328.
PR 08-NOV-2000; 2000WO-US30952.
PR 01-DEC-2000; 2000WO-US32678.

PR 20-DEC-2000; 2000WO-US34956.
PR 28-FEB-2001; 2001WO-US06520.
PR 01-JUN-2001; 2001WO-US17800.
PR 20-JUN-2001; 2001WO-US19692.
PR 29-JUN-2001; 2001WO-US21066.
PR 09-JUL-2001; 2001WO-US21735.
PR 29-AUG-2001; 2001WO-US27099.
PR 18-SEP-1997; 97US-059263P.
PR 18-SEP-1997; 97US-059266P.
PR 17-OCT-1997; 97US-062250P.
PR 21-OCT-1997; 97US-063486P.
PR 24-OCT-1997; 97US-063120P.
PR 24-OCT-1997; 97US-063121P.
PR 28-OCT-1997; 97US-063540P.
PR 28-OCT-1997; 97US-063541P.
PR 28-OCT-1997; 97US-063544P.
PR 28-OCT-1997; 97US-063564P.
PR 29-OCT-1997; 97US-063734P.
PR 31-OCT-1997; 97US-063870P.
PR 31-OCT-1997; 97US-064103P.
PR 13-NOV-1997; 97US-065311P.
PR 21-NOV-1997; 97US-066120P.
PR 24-NOV-1997; 97US-066466P.
PR 24-NOV-1997; 97US-066772P.
PR 11-DEC-1997; 97US-069335P.
PR 12-DEC-1997; 97US-069425P.
PR 17-DEC-1997; 97US-069870P.
PR 18-DEC-1997; 97US-068017P.
PR 10-MAR-1998; 98US-077450P.
PR 11-MAR-1998; 98US-077632P.
PR 11-MAR-1998; 98US-077649P.
PR 20-MAR-1998; 98US-078886P.
PR 20-MAR-1998; 98US-078939P.
PR 27-MAR-1998; 98US-079664P.
PR 27-MAR-1998; 98US-079786P.
PR 31-MAR-1998; 98US-080107P.
PR 31-MAR-1998; 98US-080194P.
PR 01-APR-1998; 98US-080327P.
PR 01-APR-1998; 98US-080333P.
PR 08-APR-1998; 98US-081049P.
PR 08-APR-1998; 98US-081070P.
PR 09-APR-1998; 98US-081195P.
PR 15-APR-1998; 98US-081838P.
PR 21-APR-1998; 98US-082568P.
PR 21-APR-1998; 98US-082569P.
PR 22-APR-1998; 98US-082704P.
PR 22-APR-1998; 98US-082797P.
PR 28-APR-1998; 98US-083322P.
PR 29-APR-1998; 98US-083495P.
PR 29-APR-1998; 98US-083496P.
PR 29-APR-1998; 98US-083499P.
PR 29-APR-1998; 98US-083559P.
PR 05-MAY-1998; 98US-084366P.
PR 06-MAY-1998; 98US-084414P.
PR 07-MAY-1998; 98US-084639P.
PR 07-MAY-1998; 98US-084640P.
PR 07-MAY-1998; 98US-084643P.
PR 15-MAY-1998; 98US-085579P.
PR 15-MAY-1998; 98US-085580P.
PR 15-MAY-1998; 98US-085582P.
PR 15-MAY-1998; 98US-085700P.
PR 18-MAY-1998; 98US-086023P.
PR 22-MAY-1998; 98US-086392P.
PR 22-MAY-1998; 98US-086486P.
PR 28-MAY-1998; 98US-087098P.
PR 28-MAY-1998; 98US-087208P.
PR 02-JUN-1998; 98US-087609P.
PR 02-JUN-1998; 98US-087759P.
PR 03-JUN-1998; 98US-087827P.
PR 04-JUN-1998; 98US-088025P.
PR 04-JUN-1998; 98US-088028P.
PR 04-JUN-1998; 98US-088029P.
PR 04-JUN-1998; 98US-088033P.

PR 04-JUN-1998; 98US-088326P.
PR 05-JUN-1998; 98US-088167P.
PR 05-JUN-1998; 98US-088202P.
PR 05-JUN-1998; 98US-088212P.
PR 05-JUN-1998; 98US-088217P.
PR 09-JUN-1998; 98US-088555P.
PR 10-JUN-1998; 98US-088722P.
PR 10-JUN-1998; 98US-088738P.
PR 10-JUN-1998; 98US-088740P.
PR 10-JUN-1998; 98US-088811P.
PR 10-JUN-1998; 98US-088824P.
PR 10-JUN-1998; 98US-088825P.
PR 10-JUN-1998; 98US-088826P.
PR 11-JUN-1998; 98US-088861P.
PR 11-JUN-1998; 98US-088863P.
PR 11-JUN-1998; 98US-088876P.
PR 12-JUN-1998; 98US-089090P.
PR 12-JUN-1998; 98US-089105P.
PR 16-JUN-1998; 98US-089512P.
PR 16-JUN-1998; 98US-089514P.
PR 17-JUN-1998; 98US-089538P.
PR 17-JUN-1998; 98US-089598P.
PR 17-JUN-1998; 98US-089653P.
PR 18-JUN-1998; 98US-089908P.
PR 19-JUN-1998; 98US-089952P.
PR 22-JUN-1998; 98US-090246P.
PR 22-JUN-1998; 98US-090252P.
PR 22-JUN-1998; 98US-090254P.
PR 24-JUN-1998; 98US-090429P.
PR 24-JUN-1998; 98US-090435P.
PR 24-JUN-1998; 98US-090444P.
PR 24-JUN-1998; 98US-090461P.
PR 24-JUN-1998; 98US-090535P.
PR 24-JUN-1998; 98US-090540P.
PR 25-JUN-1998; 98US-090676P.
PR 25-JUN-1998; 98US-090678P.
PR 25-JUN-1998; 98US-090688P.
PR 25-JUN-1998; 98US-090690P.
PR 25-JUN-1998; 98US-090694P.
PR 25-JUN-1998; 98US-090695P.
PR 25-JUN-1998; 98US-090696P.
PR 26-JUN-1998; 98US-090862P.
PR 26-JUN-1998; 98US-090863P.
PR 26-JUN-1998; 98US-091010P.
PR 01-JUL-1998; 98US-091359P.
PR 01-JUL-1998; 98US-091544P.
PR 02-JUL-1998; 98US-091478P.
PR 02-JUL-1998; 98US-091486P.
PR 02-JUL-1998; 98US-091626P.
PR 02-JUL-1998; 98US-091628P.
PR 02-JUL-1998; 98US-091632P.
PR 24-JUL-1998; 98US-094006P.
PR 04-AUG-1998; 98US-095282P.
PR 10-AUG-1998; 98US-095998P.
PR 10-AUG-1998; 98US-096012P.
PR 17-AUG-1998; 98US-096757P.
PR 17-AUG-1998; 98US-096766P.
PR 17-AUG-1998; 98US-096867P.
PR 17-AUG-1998; 98US-096891P.
PR 17-AUG-1998; 98US-096897P.
PR 18-AUG-1998; 98US-096949P.
PR 18-AUG-1998; 98US-096959P.
PR 18-AUG-1998; 98US-097022P.
PR 26-AUG-1998; 98US-097952P.
PR 26-AUG-1998; 98US-097954P.
PR 26-AUG-1998; 98US-097955P.
PR 26-AUG-1998; 98US-097971P.
PR 26-AUG-1998; 98US-097974P.
PR 26-AUG-1998; 98US-098014P.
PR 01-SEP-1998; 98US-098716P.
PR 01-SEP-1998; 98US-098723P.
PR 02-SEP-1998; 98US-098803P.
PR 02-SEP-1998; 98US-098821P.

```
PR 02-SEP-1998; 98US-098843P.
PR 09-SEP-1998; 98US-099602P.
PR 10-SEP-1998; 98US-099741P.

Query Match      98.7%; Score 1092; DB 24; Length 382;
Best Local Similarity 96.2%; Pred. No. 7e-103;
Matches 204; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 GRLLS-----GQVCRGGTQPCYKVIYFHTSRRLNFEFAKEACRRDGGQLVSIES 52
Db 24 GRLLSASDLDRGGQVCRGGTQPCYKVIYFHTSRRLNFEFAKEACRRDGGQLVSIES 83

QY 53 EDEQKLIKFIENLLPSDGFWIGLRRRREKQSNSTACQDLYAWTDGSIQFRNMYVDEP 112
Db 84 EDEQKLIKFIENLLPSDGFWIGLRRRREKQSNSTACQDLYAWTDGSIQFRNMYVDEP 143

QY 113 SCGSEVCVMYHQPSAPAGIGGYPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGEET 172
Db 144 SCGSEVCVMYHQPSAPAGIGGYPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGEET 203

QY 173 ELTTPVLPEETQEEDAKKTFKESREAAALNLAY 204
Db 204 ELTTPVLPEETQEEDAKKTFKESREAAALNLAY 235

RESULT 15
ABU65911
ID ABU65911 standard; Protein; 382 AA.
XX
AC ABU65911;
XX
DT 20-MAY-2003 (first entry)
XX
DE Novel human secreted and transmembrane protein PRO234.
XX
KW Human; secreted protein; transmembrane protein; cytostatic;
KW gene Therapy; TNF-Agonist-Alpha; chondrocyte stimulator; tumour;
KW adrenal tumour; lung tumour; colon tumour; breast tumour;
KW prostate tumour; rectal tumour; cervical tumour; liver tumour.
XX
OS Homo sapiens.
XX
PN US2003036157-A1.
XX
PD 20-FEB-2003.
XX
PF 02-JUL-2002; 2002US-0188769.
XX
PR 16-SEP-1998; 98WO-US19330.
PR 07-OCT-1998; 98WO-US21141.
PR 01-DEC-1998; 98WO-US25108.
PR 08-MAR-1999; 99WO-US05028.
PR 14-MAY-1999; 99WO-US10733.
PR 02-JUN-1999; 99WO-US12252.
PR 01-SEP-1999; 99WO-US20111.
PR 15-SEP-1999; 99WO-US21090.
PR 01-DEC-1999; 99WO-US28301.
PR 02-DEC-1999; 99WO-US28551.
PR 30-DEC-1999; 99WO-US31274.
PR 05-JAN-2000; 2000WO-US00219.
PR 18-FEB-2000; 2000WO-US04341.
PR 18-FEB-2000; 2000WO-US04342.
PR 22-FEB-2000; 2000WO-US04141.
PR 24-FEB-2000; 2000WO-US05004.
PR 01-MAR-2000; 2000WO-US05601.
PR 02-MAR-2000; 2000WO-US05841.
PR 15-MAR-2000; 2000WO-US06884.
PR 30-MAR-2000; 2000WO-US08439.
PR 17-MAY-2000; 2000WO-US13705.
PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.
PR 02-JUN-2000; 2000WO-US15264.
PR 28-JUL-2000; 2000WO-US20710.

PR 24-AUG-2000; 2000WO-US23328.
PR 08-NOV-2000; 2000WO-US30952.
PR 01-DEC-2000; 2000WO-US32678.
PR 20-DEC-2000; 2000WO-US34956.
PR 28-FEB-2001; 2001WO-US06520.
PR 01-JUN-2001; 2001WO-US17800.
PR 20-JUN-2001; 2001WO-US19692.
PR 29-JUN-2001; 2001WO-US21066.
PR 09-JUL-2001; 2001WO-US21735.
PR 29-AUG-2001; 2001WO-US27099.
PR 18-SEP-1997; 97US-059263P.
PR 18-SEP-1997; 97US-059266P.
PR 17-OCT-1997; 97US-062250P.
PR 21-OCT-1997; 97US-063486P.
PR 24-OCT-1997; 97US-063120P.
PR 24-OCT-1997; 97US-063121P.
PR 28-OCT-1997; 97US-063540P.
PR 28-OCT-1997; 97US-063541P.
PR 28-OCT-1997; 97US-063544P.
PR 28-OCT-1997; 97US-063564P.
PR 29-OCT-1997; 97US-063734P.
PR 31-OCT-1997; 97US-063870P.
PR 31-OCT-1997; 97US-064103P.
PR 13-NOV-1997; 97US-065311P.
PR 21-NOV-1997; 97US-066120P.
PR 24-NOV-1997; 97US-066466P.
PR 24-NOV-1997; 97US-066772P.
PR 11-DEC-1997; 97US-069335P.
PR 12-DEC-1997; 97US-069425P.
PR 17-DEC-1997; 97US-069870P.
PR 18-DEC-1997; 97US-068017P.
PR 10-MAR-1998; 98US-077450P.
PR 11-MAR-1998; 98US-077632P.
PR 11-MAR-1998; 98US-077649P.
PR 20-MAR-1998; 98US-078886P.
PR 20-MAR-1998; 98US-078939P.
PR 27-MAR-1998; 98US-079664P.
PR 27-MAR-1998; 98US-079786P.
PR 31-MAR-1998; 98US-080107P.
PR 31-MAR-1998; 98US-080194P.
PR 01-APR-1998; 98US-080327P.
PR 01-APR-1998; 98US-080333P.
PR 08-APR-1998; 98US-081049P.
PR 08-APR-1998; 98US-081070P.
PR 09-APR-1998; 98US-081195P.
PR 15-APR-1998; 98US-081838P.
PR 21-APR-1998; 98US-082568P.
PR 21-APR-1998; 98US-082569P.
PR 22-APR-1998; 98US-082704P.
PR 22-APR-1998; 98US-082797P.
PR 28-APR-1998; 98US-083322P.
PR 29-APR-1998; 98US-083495P.
PR 29-APR-1998; 98US-083496P.
PR 29-APR-1998; 98US-083499P.
PR 29-APR-1998; 98US-083559P.
PR 05-MAY-1998; 98US-084366P.
PR 06-MAY-1998; 98US-084414P.
PR 07-MAY-1998; 98US-084639P.
PR 07-MAY-1998; 98US-084640P.
PR 07-MAY-1998; 98US-084643P.
PR 15-MAY-1998; 98US-085579P.
PR 15-MAY-1998; 98US-085580P.
PR 15-MAY-1998; 98US-085582P.
PR 15-MAY-1998; 98US-085700P.
PR 18-MAY-1998; 98US-086023P.
PR 22-MAY-1998; 98US-086392P.
PR 22-MAY-1998; 98US-086486P.
PR 28-MAY-1998; 98US-087098P.
PR 28-MAY-1998; 98US-087208P.
PR 02-JUN-1998; 98US-087609P.
PR 02-JUN-1998; 98US-087759P.
PR 03-JUN-1998; 98US-087827P.
PR 04-JUN-1998; 98US-088025P.
```

PR 04-JUN-1998; 98US-088028P.
PR 04-JUN-1998; 98US-088029P.
PR 04-JUN-1998; 98US-088033P.
PR 04-JUN-1998; 98US-088326P.
PR 05-JUN-1998; 98US-088167P.
PR 05-JUN-1998; 98US-088202P.
PR 05-JUN-1998; 98US-088212P.
PR 05-JUN-1998; 98US-088217P.
PR 09-JUN-1998; 98US-088655P.
PR 10-JUN-1998; 98US-088722P.
PR 10-JUN-1998; 98US-088738P.
PR 10-JUN-1998; 98US-088740P.
PR 10-JUN-1998; 98US-088811P.
PR 10-JUN-1998; 98US-088824P.
PR 10-JUN-1998; 98US-088825P.
PR 10-JUN-1998; 98US-088826P.
PR 11-JUN-1998; 98US-088861P.
PR 11-JUN-1998; 98US-088863P.
PR 11-JUN-1998; 98US-088876P.
PR 12-JUN-1998; 98US-089090P.
PR 12-JUN-1998; 98US-089105P.
PR 16-JUN-1998; 98US-089512P.
PR 16-JUN-1998; 98US-089514P.
PR 17-JUN-1998; 98US-089538P.
PR 17-JUN-1998; 98US-089598P.
PR 17-JUN-1998; 98US-089653P.
PR 18-JUN-1998; 98US-089908P.
PR 19-JUN-1998; 98US-089952P.
PR 22-JUN-1998; 98US-090246P.
PR 22-JUN-1998; 98US-090252P.
PR 22-JUN-1998; 98US-090254P.
PR 24-JUN-1998; 98US-090429P.
PR 24-JUN-1998; 98US-090435P.
PR 24-JUN-1998; 98US-090444P.
PR 24-JUN-1998; 98US-090461P.
PR 24-JUN-1998; 98US-090535P.
PR 24-JUN-1998; 98US-090540P.
PR 25-JUN-1998; 98US-090676P.
PR 25-JUN-1998; 98US-090678P.
PR 25-JUN-1998; 98US-090688P.
PR 25-JUN-1998; 98US-090690P.
PR 25-JUN-1998; 98US-090694P.
PR 25-JUN-1998; 98US-090695P.
PR 26-JUN-1998; 98US-090696P.
PR 26-JUN-1998; 98US-090862P.
PR 26-JUN-1998; 98US-090863P.
PR 26-JUN-1998; 98US-091010P.
PR 01-JUL-1998; 98US-091359P.
PR 01-JUL-1998; 98US-091544P.
PR 02-JUL-1998; 98US-091478P.
PR 02-JUL-1998; 98US-091486P.
PR 02-JUL-1998; 98US-091626P.
PR 02-JUL-1998; 98US-091628P.
PR 02-JUL-1998; 98US-091632P.
PR 24-JUL-1998; 98US-094006P.
PR 04-AUG-1998; 98US-095282P.
PR 10-AUG-1998; 98US-095998P.
PR 10-AUG-1998; 98US-096012P.
PR 17-AUG-1998; 98US-096757P.
PR 17-AUG-1998; 98US-096766P.
PR 17-AUG-1998; 98US-096867P.
PR 17-AUG-1998; 98US-096891P.
PR 17-AUG-1998; 98US-096897P.
PR 18-AUG-1998; 98US-096949P.
PR 18-AUG-1998; 98US-096959P.
PR 18-AUG-1998; 98US-097022P.
PR 26-AUG-1998; 98US-097952P.
PR 26-AUG-1998; 98US-097954P.
PR 26-AUG-1998; 98US-097955P.
PR 26-AUG-1998; 98US-097971P.
PR 26-AUG-1998; 98US-097974P.
PR 26-AUG-1998; 98US-098014P.
PR 01-SEP-1998; 98US-098716P.

PR 01-SEP-1998; 98US-098723P.
PR 02-SEP-1998; 98US-098803P.
PR 02-SEP-1998; 98US-098821P.
PR 02-SEP-1998; 98US-098843P.
PR 09-SEP-1998; 98US-099602P.
PR 10-SEP-1998; 98US-099741P.
PR 10-SEP-1998; 98US-099754P.
PR 10-SEP-1998; 98US-099763P.

Query Match 98.7%; Score 1092; DB 24; Length 382;
Best Local Similarity 96.2%; Pred. No. 7e-103;
Matches 204; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

Qy 1 GRLLS-----GQPVCRGGTQRPCYKVIYFHDTSRRLNFEFEAKEACRRDGGQLVSIES 52
|||||
Db 24 GRLLSASDDLRLGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEFEAKEACRRDGGQLVSIES 83
|||||

Qy 53 EDEQKLIKFIENLLPSDGFWIGLRRREEKQSNSTACQDLYAWTDGSI SQFRNWWYVDEP 112
|||||

Db 84 EDEQKLIKFIENLLPSDGFWIGLRRREEKQSNSTACQDLYAWTDGSI SQFRNWWYVDEP 143
|||||

Qy 113 SCGSEVCVVMYHQPSAPAGIGGYPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGET 172
|||||

Db 144 SCGSEVCVVMYHQPSAPAGIGGYPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGET 203
|||||

Qy 173 ELTTPVLPEETOEDAKKTFKESREAAALNLAY 204
|||||

Db 204 ELTTPVLPEETOEDAKKTFKESREAAALNLAY 235
|||||

Search completed: December 22, 2003, 16:10:08
Job time : 29.8986 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 22, 2003, 16:04:13 ; Search time 28.7569 Seconds
(without alignments)
1120.478 Million cell updates/sec

Title: US-09-887-855-2_COPY_25_227
Perfect score: 1100
Sequence: 1 RLISGQPVCRGQTQPCYKV.....EEDAKTKFKESREAAALNLAY 203

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 20000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq 19Jun03.*
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1100	100.0	374	21	AAY93948
2	1100	100.0	374	21	AAE03651
3	1100	100.0	374	23	ABG66680
4	1100	100.0	374	23	ABB90203
5	1100	100.0	387	22	AAM25796
6	1090	99.1	374	21	AA91490
7	1086	98.7	382	20	AA913367
8	1086	98.7	382	22	AAU29033
9	1086	98.7	382	22	AAB80235

10	1086	98.7	382	24	ABU69645	Novel human secret
11	1086	98.7	382	24	ABU71121	Human PRO234 prote
12	1086	98.7	382	24	ABU71468	Human PRO polypept
13	1086	98.7	382	24	ABU71914	Human secreted/tra
14	1086	98.7	382	24	ABU65578	Human secreted/tra
15	1086	98.7	382	24	ABU65911	Novel human secret
16	1086	98.7	382	24	ABU67368	Human secreted pro
17	1086	98.7	382	24	ABU67415	Human secreted/tra
18	1086	98.7	382	24	ABU64522	Human secreted/tra
19	1086	98.7	382	24	ABU65273	Human PRO polypept
20	1086	98.7	382	24	ABU58409	Human PRO polypept
21	1086	98.7	382	24	ABU55945	Human secreted/tra
22	1086	98.7	382	24	ABU56940	Human PRO polypept
23	1086	98.7	382	24	ABU54370	Human secreted/tra
24	1086	98.7	382	24	ABU10519	Human secreted/tra
25	561.5	51.0	273	21	AA18913	A novel polypeptid
26	561.5	51.0	273	22	AAU12441	Human PRO1890 poly
27	561.5	51.0	273	22	AA187309	Human C-type lecti
28	561.5	51.0	273	22	AA187609	Human PRO1890. Ho
29	561.5	51.0	273	23	ABG95934	Human secreted/tra
30	561.5	51.0	273	23	ABB95554	Human angiogenesis
31	561.5	51.0	273	23	ABB84948	Human PRO1890 prot
32	561.5	51.0	273	24	ABU69084	Human PRO polypept
33	561.5	51.0	273	24	ABU69107	Human PRO polypept
34	561.5	51.0	273	24	ABU71589	Human secreted pol
35	561.5	51.0	273	24	ABU72035	Novel human secret
36	561.5	51.0	273	24	ABU72192	Human PRO polypept
37	561.5	51.0	273	24	ABU66839	Human PRO polypept
38	561.5	51.0	273	24	ABU67115	Human secreted/tra
39	561.5	51.0	273	24	ABU59920	Novel secreted and
40	534.5	48.6	274	23	ABP69211	Human polypeptide
41	504.5	45.9	232	22	ABG94192	Human protein sequ
42	275	25.0	102	22	ABG51278	Human liver peptid
43	275	25.0	102	22	ABB21788	Protein #3787 enco
44	275	25.0	102	22	AAM29941	Peptide #3978 enco
45	225	20.5	81	21	AAY91643	Human secreted pro

ALIGNMENTS

RESULT 1
AAY93948
ID AAY93948 standard; Protein; 374 AA.
XX
AC AAY93948;
XX
DT 03-OCT-2000 (first entry)
XX
DE Amino acid sequence of a lectin ss3939 polypeptide.
XX
KW Human; lectin ss3939; chromosome 11; gene therapy.
XX
OS Homo sapiens.

Key	Location/Qualifiers
FT Peptide	1..21
FT Domain	/note= "signal peptide"
FT Domain	22..227
FT Domain	/note= "extracellular coding region"
FT Domain	228..248
FT Domain	/note= "predicted transmembrane domain"
FT Domain	249..374
FT	/note= "predicted cytoplasmic or intracellular domain"

WO200039296-A1.

06-JUL-2000.

22-DEC-1999; 99WO-US30523.

23-DEC-1998; 98US-0113820.

PA (IMMV) IMMUNEX CORP.
XX
PI Anderson DA;
XX
DR WPI; 2000-452394/39.
DR N-PSDB; AAA57382.
XX
PT ss3939 nucleic acids, polypeptides and antibodies, useful for
PT identifying human chromosome 11 and diseases associated with it -
XX
XX
PS Claim 12; Page 8; 73pp; English.
XX
CC The present sequence represents a human lectin ss3939 polypeptide. The
CC polynucleotide sequence is a source of probes, which may be used
CC to identify nucleic acids encoding ss3939 proteins, to identify human
CC chromosome number 11, to map genes on human chromosome number 11, to
CC identify diseases associated with chromosome 11, as single-stranded
CC sense or antisense oligonucleotides to inhibit expression of
CC polypeptides encoded by the ss3939 gene, and for gene therapy. The
CC ss3939 polypeptides may be useful for developing treatments for
CC diseases (none specified) associated with defective or insufficient
CC amounts of the polypeptides. The antibodies may be useful for
CC detecting the presence of ss3939 polypeptides.
XX
SQ Sequence 374 AA;

Query Match 100.0%; Score 1100; DB 21; Length 374;
Best Local Similarity 100.0%; Pred. No. 1e-103;
Matches 203; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLLSGQPVCRGGTQPCYKVIYFHDTSRRLNFEAEAKRRDGGQLVSI ESEDEQKLI EK 60
Db 25 RLLSGQPVCRGGTQPCYKVIYFHDTSRRLNFEAEAKRRDGGQLVSI ESEDEQKLI EK 84
QY 61 FIENLLPSDGDWIGLRRREEKQSNSTACODLYAWTDGSGISQFRNWWYDEPSCGSEVCV 120
Db 85 FIENLLPSDGDWIGLRRREEKQSNSTACODLYAWTDGSGISQFRNWWYDEPSCGSEVCV 144
QY 121 MYHQPSAPAGIGGPFYMFQWNDRCNMKNFICKYSDEKPAVPSREAEGETELTTPVLPE 180
Db 145 MYHQPSAPAGIGGPFYMFQWNDRCNMKNFICKYSDEKPAVPSREAEGETELTTPVLPE 204
QY 181 ETQEEDAKKTFKESREAAINLAY 203
Db 205 ETQEEDAKKTFKESREAAINLAY 227

RESULT 2
AAE03651
ID AAE03651 standard; Protein; 374 AA.
XX
AC AAE03651;
XX
DT 06-AUG-2001 (first entry)
XX
DE Human extracellular matrix and cell adhesion molecule-15 (XMAD-15).
XX
KW Human; extracellular matrix and cell adhesion molecule; XMAD;
KW gene therapy; genetic disorder; adrenoleukodystrophy; leukaemia;
KW Down's syndrome; cystic fibrosis; Gaucher's disease; myotonic dystrophy;
KW sickle cell anaemia; thalassaemia; autoimmune disorder; adenocarcinoma;
KW inflammatory disorder; acquired immune deficiency syndrome; AIDS;
KW Addison's disease; allergy; anaemia; asthma; atherosclerosis; melanoma;
KW Crohn's disease; diabetes mellitus; atopic dermatitis; lymphoma; cancer;
KW glomerulonephritis; multiple sclerosis; Grave's disease; osteoarthritis;
KW osteoporosis; psoriasis; rheumatoid arthritis; ulcerative colitis;
KW infection; cell proliferative disorder; actinic keratosis; myeloma;
KW arteriosclerosis; neutropic; anticonvulsant; antithyroid; nephrotropic;
KW neuroprotective; dermatological.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers

FT Peptide 1..24
FT /label= Signal_peptide
FT 25..374
FT /note= "Mature human extracellular matrix and cell
FT adhesion molecule (XMAD)"
FT 46..63
FT /note= "C-type lectin domain"
FT 163..176
FT /note= "C-type lectin domain"
FT 224..247
FT /note= "Transmembrane motif"
FT 328..348
FT /note= "Transmembrane motif"
FT
XX WO200142285-A2.
PN 14-JUN-2001.
XX
PD
XX
PF 05-DEC-2000; 2000WO-US32990.
XX
PR 10-DEC-1999; 99US-0172852.
PR 16-DEC-1999; 99US-0172354.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Yue H, Tang YT, Lal P, Burford N, Azimzai Y, Patterson C;
PI Baughn MR, Lu DAM, Shah P, Au-Young J;
XX
DR WPI; 2001-381632/40.
DR N-PSDB; AAD08059.
XX
PT New human extracellular matrix and cell adhesion molecules and
PT polynucleotide sequences encoding them, useful for diagnosis,
PT prevention, treatment of genetic, autoimmune and cell proliferative
PT disorders
XX
PS Claim 1; Page 108-109; 135pp; English.
XX
CC The present sequence is a human extracellular matrix and cell
CC adhesion molecule (XMAD). The XMAD is used for screening a compound for
CC effectiveness as an agonist or antagonist of XMAD. The identified agonist
CC or antagonist are used for treating a disease or condition associated
CC with decreased or increased expression of functional XMAD. The
CC polynucleotides encoding XMAD are useful in somatic or germline gene
CC therapy to correct a genetic deficiency, to express a conditionally
CC lethal gene product and to express a protein which affords protection
CC against intracellular parasites and also for diagnosis of disorders
CC associated with expression of XMAD. They are also used for generating
CC hybridisation probes useful in mapping the naturally occurring genomic
CC sequences and to create knock in humanised animals (pigs) or transgenic
CC animals (mice or rats) to model human diseases. Oligonucleotide or longer
CC fragments derived from the polynucleotide sequences may be used as
CC elements on a microarray. Antibodies which specifically bind XMAD may be
CC used for the diagnosis of disorders associated with the expression of
CC XMAD, or in assays to monitor patients being treated with XMAD. Diseases
CC diagnosed, prevented or treated include genetic disorders such as
CC adrenoleukodystrophy, Down's syndrome, cystic fibrosis, Gaucher's
CC disease, myotonic dystrophy, sickle cell anaemia, thalassaemia,
CC autoimmune/inflammatory disorders such as acquired immune deficiency
CC syndrome (AIDS), Addison's disease, allergies, anaemia, asthma,
CC atherosclerosis, Crohn's disease, diabetes mellitus, atopic dermatitis,
CC glomerulonephritis, multiple sclerosis, Grave's disease, osteoarthritis,
CC osteoporosis, psoriasis, rheumatoid arthritis, ulcerative colitis,
CC bacterial, fungal, parasitic, protozoal and helminthic infections and
CC cell proliferative disorders such as actinic keratosis, arteriosclerosis
CC and cancer including breast, bladder, bone marrow, brain and uterus
CC cancer, leukaemia, adenocarcinoma, lymphoma, melanoma and myeloma.
XX
SQ Sequence 374 AA;

Query Match 100.0%; Score 1100; DB 22; Length 374;
Best Local Similarity 100.0%; Pred. No. 1e-103;
Matches 203; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLLSGQPVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIIESEDEQKLIIEK 60
DB 25 RLLSGQPVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIIESEDEQKLIIEK 84
QY 61 FIENLLPSDGDWIGLRRREKQSNSTACQDLYAWTDGSIQFRNWWYVDEPSCGSEVCVW 120
DB 85 FIENLLPSDGDWIGLRRREKQSNSTACQDLYAWTDGSIQFRNWWYVDEPSCGSEVCVW 144
QY 121 MYHQSAPAGIGGPFYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGETELTTPVLPE 180
DB 145 MYHQSAPAGIGGPFYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGETELTTPVLPE 204
QY 181 ETQEEDAKKTFKESREAAALNLAY 203
DB 205 ETQEEDAKKTFKESREAAALNLAY 227

RESULT 3

ABG66680
ID ABG66680 standard; Protein; 374 AA.

XX AC ABG66680;

XX 30-AUG-2002 (first entry)

XX Human novel polypeptide #15.

KW Human; inflammatory condition; shock; sepsis; immune response;
KW cancer; wound healing; central nervous system disease; haematopoiesis;
KW peripheral nervous system disease; amyotrophic lateral sclerosis; tendon;
KW myeloid cell disorder; lymphoid cell disorder; platelet disorder; bone;
KW cartilage; ligament; nerve tissue; ulcer; osteoporosis; osteoarthritis;
KW bone degenerative disorder; periodontal disease; reperfusion injury;
KW lung fibrosis; liver fibrosis; autoimmune disorder; bacterial infection;
KW allergic condition; thrombolysis; thrombosis; coagulation disorder;
KW fungal infection.

XX OS Homo sapiens.

XX PN WO200244340-A2.

XX 06-JUN-2002.

XX 30-NOV-2001; 2001WO-US47004.

XX 30-NOV-2000; 2000US-0028952.

XX (HYSE-) HYSEQ INC.

PI Tang YT, Goodrich RW, Liu C, Zhou P, Asundi V, Wang J, Wang D;
PI Yamazaki V, Ujwal ML, Drmanac RT;

XX WPI; 2002-508509/54.

XX N-PSDB; ABK94904.

XX Novel nucleic acids and polypeptides for diagnosis, treatment of
PT inflammatory, autoimmune, nervous system, myeloid or lymphoid cell
PT disorders, cancer and promoting wound healing -

PS Claim 10; Page 579-580; 672pp; English.

XX The invention relates to human novel polynucleotides and associated
CC polypeptides. The polynucleotides and polypeptides are useful for
CC treating inflammatory conditions such as arthritis, nephritis, Crohn's
CC disease, ischaemia-reperfusion injury, shock, sepsis, immune responses
CC and cancer and for promoting wound healing. The sequences are used to
CC induce the proliferation of neural cells and regeneration of nerve and
CC brain tissue, and are useful for the treatment of central and peripheral
CC nervous system diseases and neuropathies, such as Alzheimer's disease,
CC Parkinson's disease, Huntington's disease and amyotrophic lateral
CC sclerosis. The sequences are involved in chemotactic or chemokinetic
CC activity, regulation of haematopoiesis, treatment of myeloid or lymphoid

CC cell disorders and platelet disorders such as thrombocytopenia,
CC regeneration of bone, cartilage, tendon, ligament and/or nerve tissue
CC growth, tissue repair, healing of burns, incisions, ulcers, treatment of
CC osteoporosis, osteoarthritis, bone degenerative disorders and periodontal
CC disease. The sequences of the invention are also useful for gut
CC protection or regeneration and treatment of lung or liver fibrosis,
CC reperfusion injury in various tissues, immune deficiencies and disorders
CC including severe combined immunodeficiency (SCID), bacterial or fungal
CC infections, autoimmune disorders e.g. multiple sclerosis and myasthenia
CC gravis, allergic conditions such as asthma, thrombolysis or thrombosis
CC and coagulation disorders. Sequences ABG66666-ABG66758 represent human
CC novel polypeptides of the invention.

XX SQ Sequence 374 AA;

Query Match 100.0%; Score 1100; DB 23; Length 374;
Best Local Similarity 100.0%; Pred. NO. 1e-103;
Matches 203; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLLSGQPVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIIESEDEQKLIIEK 60
DB 25 RLLSGQPVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIIESEDEQKLIIEK 84
QY 61 FIENLLPSDGDWIGLRRREKQSNSTACQDLYAWTDGSIQFRNWWYVDEPSCGSEVCVW 120
DB 85 FIENLLPSDGDWIGLRRREKQSNSTACQDLYAWTDGSIQFRNWWYVDEPSCGSEVCVW 144
QY 121 MYHQSAPAGIGGPFYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGETELTTPVLPE 180
DB 145 MYHQSAPAGIGGPFYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGETELTTPVLPE 204
QY 181 ETQEEDAKKTFKESREAAALNLAY 203
DB 205 ETQEEDAKKTFKESREAAALNLAY 227

RESULT 4

ABB90203

ID ABB90203 standard; Protein; 374 AA.

XX AC ABB90203;

XX 24-MAY-2002 (first entry)

XX Human polypeptide SEQ ID NO 2579.

DE Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein.

OS Homo sapiens.

XX WO200190304-A2.

XX 29-NOV-2001.

XX 18-MAY-2001; 2001WO-US16450.

XX 19-MAY-2000; 2000US-205515P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Birse CE, Rosen CA;

XX WPI; 2002-122018/16.

XX N-PSDB; ABL90612.

XX Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
PT prevention of neural, immune system, muscular, reproductive,
PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
PT disorders -

XX PS Claim 11; SEQ ID NO 2579; 2081pp + Sequence Listing; English.

XX CC The invention relates to novel genes (ABL89449-ABL90853) and proteins (ABB89040-ABB90444) useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. The genes are isolated from a range of human tissues disclosed in the specification.

CC CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver, lung, or urogenital;

CC CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders such as myocardial ischaemias; (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f) infectious diseases such as viral, bacterial, fungal and parasitic infections.

CC CC Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 374 AA;

Query Match 100.0%; Score 1100; DB 23; Length 374;

Best Local Similarity 100.0%; Pred. No. 1e-103;

Matches 203; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLLSGQPVCRGGTQPCYKVIYFHDTSRRLLNFEEAKEACRRDGGQLVSI ESEDEQKLIK 60

Db 25 RLLSGQPVCRGGTQPCYKVIYFHDTSRRLLNFEEAKEACRRDGGQLVSI ESEDEQKLIK 84

QY 61 FIENLLPSDGDGFWIGLRRREEKQSNSTACQDLYAWTDGSI SQFRNWWYVDEPSCGSEVCV 120

Db 85 FIENLLPSDGDGFWIGLRRREEKQSNSTACQDLYAWTDGSI SQFRNWWYVDEPSCGSEVCV 144

QY 121 MYHQPSAPAGIGGYPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGETELTTPVLPE 180

Db 145 MYHQPSAPAGIGGYPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGETELTTPVLPE 204

QY 181 ETQEEDAKKTFKESREAAALNLAY 203

Db 205 ETQEEDAKKTFKESREAAALNLAY 227

RESULT 5

AAM25796

ID AAM25796 standard; Protein; 387 AA.

XX AC AAM25796;

XX 16-OCT-2001 (first entry)

DE Human protein sequence SEQ ID NO:1311.

XX KW Human; cancer; ulcer; HIV infection; human immunodeficiency virus; antiinflammatory; antirheumatic; antiarthritic; immunosuppressive; antibacterial; endocrine; cardiant; central nervous system; virucide; anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia; antiaggagant; haemostatic; vulnery; antiulcer; osteopathic; eczema; dermatological; antiallergic; antiasthmatic; antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic; antiparkinsonian; infection; immunostimulant; gene therapy; antisense therapy; vaccine; inflammation; antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis; cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity; genetic disease; haematopoietic disorder; platelet disorder; asthma; thrombocytopaenia; osteoporosis; severe combined immunodeficiency; allergic rhinitis; diabetes; multiple sclerosis; depression; Alzheimer's disease; Parkinson's disease; neurodegenerative disorder; neurological disorder.

OS Homo sapiens.

XX

PN WO200153455-A2.

XX 26-JUL-2001.

PD

XX 22-DEC-2000; 2000WO-US35017.

PF

XX 23-DEC-1999; 99US-0471275.

PR

XX 21-JAN-2000; 2000US-0488725.

PR

XX 25-APR-2000; 2000US-0552317.

PR

XX (HYSE-) HYSEQ INC.

PA

XX Tang YT, Liu C, Drmanac RT;

PI

XX WPI; 2001-457603/49.

DR

XX N-PSDB; AAH99737.

DR

XX Isolated human polynucleotides encoding polypeptides, useful for the treatment and diagnosis of e.g. cancer, ulcers and HIV infection -

PT

PT Claim 20; Page 272; 1217pp; English.

PS

XX

XX AAH99166 to AAH99904 encode the human proteins given in AAM25225 to AAM25963. The proteins can have activities based on the tissues and cells they are expressed in, such as: antiinflammatory; antirheumatic; antiarthritic; immunosuppressive; antibacterial; endocrine; cardiant; central nervous system; virucide; anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; antiaggagant; haemostatic; vulnery; antiulcer; osteopathic; dermatological; antiallergic; antidepressant; nootropic; antidiabetic; cytostatic; neuroprotective; antiasthmatic; antidiabetic; antiparkinsonian; and immunostimulant. The proteins and polynucleotides encoding them can be used in gene therapy, antisense therapy and vaccine production. The proteins and polynucleotides are useful for screening for agonists or antagonists of a protein and for the treatment and diagnosis of disorders associated with the activity of a protein e.g. inflammation, rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction, neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal infections, autoimmunity, genetic diseases, haematopoietic disorders, anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers, osteoporosis, severe combined immunodeficiency, eczema, allergic rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression, Alzheimer's disease, Parkinson's disease, neurodegenerative and neurological disorders.

XX SQ Sequence 387 AA;

Query Match 100.0%; Score 1100; DB 22; Length 387;

Best Local Similarity 100.0%; Pred. No. 1.e-103;

Matches 203; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLLSGQPVCRGGTQPCYKVIYFHDTSRRLLNFEEAKEACRRDGGQLVSI ESEDEQKLIK 60

Db 38 RLLSGQPVCRGGTQPCYKVIYFHDTSRRLLNFEEAKEACRRDGGQLVSI ESEDEQKLIK 97

QY 61 FIENLLPSDGDGFWIGLRRREEKQSNSTACQDLYAWTDGSI SQFRNWWYVDEPSCGSEVCV 120

Db 98 FIENLLPSDGDGFWIGLRRREEKQSNSTACQDLYAWTDGSI SQFRNWWYVDEPSCGSEVCV 157

QY 121 MYHQPSAPAGIGGYPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGETELTTPVLPE 180

Db 158 MYHQPSAPAGIGGYPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGETELTTPVLPE 217

QY 181 ETQEEDAKKTFKESREAAALNLAY 203

Db 218 ETQEEDAKKTFKESREAAALNLAY 240

RESULT 6

AAAY91490

ID AAAY91490 standard; Protein; 374 AA.

XX

AC AAAY91490;

XX

DT 29-JUN-2000 (first entry)

XX Human secreted protein sequence encoded by gene 40 SEQ ID NO:163.

DE Human; secreted protein; diagnosis; cytostatic; immunosuppressive;

XX antiHIV; antiinflammatory; nontropic; neuroprotective; antiallergic;

KW osteopathic; antiarthritic; antibacterial; antidiabetic; antiasthma;

KW antipsoriatic; cardiant; gene therapy; cancer; neurological disorder;

KW immune disease; inflammation; blood disorder; tumour.

XX Homo sapiens.

OS WO200006698-A1.

XX 10-FEB-2000.

PN 29-JUL-1999; 99WO-US17130.

XX 30-JUL-1998; 98US-0094657.

PR 05-AUG-1998; 98US-0095486.

PR 06-AUG-1998; 98US-0095454.

PR 06-AUG-1998; 98US-0095455.

PR 12-AUG-1998; 98US-0096319.

XX (HUMA-) HUMAN GENOME SCI INC.

PA Komatsoulis GA, Rosen CA, Ruben SM, Duan R, Moore PA, Shi Y;

XX Lafleur D, Wei Y, Ni J, Florence KA, Young PE, Brewer LA;

PI Soppet DR, Endress GA, Ebner R, Olsen HS, Mucenski M;

XX WPI; 2000-195282/17.

DR N-PSDB; AAA26385.

DR New isolated human genes and the secreted polypeptides they encode,

XX useful for diagnosis and treatment of e.g. cancers, neurological

PT disorders, immune diseases, inflammation or blood disorders -

PT Claim 11; Page 483-484; 634pp; English.

XX The polynucleotide sequences given in AAA26346 to AAA26458 encode the

CC human secreted proteins given in AAY91451 to AAY91691. The human secreted

CC proteins can have activities based on the tissues and cells they are

CC expressed in. Examples of the activities are: cytostatic;

CC immunosuppressive; antiHIV; antiinflammatory; nontropic; neuroprotective;

CC antiallergic; osteopathic; antiarthritic; antibacterial; antidiabetic;

CC antiasthma; antipsoriatic; and cardiant. The polynucleotides and their

CC corresponding secreted proteins are useful for preventing, treating or

CC ameliorating medical conditions, e.g. by protein or gene therapy. Also

CC pathological conditions can be diagnosed by determining the amount of the

CC proteins in a sample or by determining the presence of mutations in the

CC polynucleotides. Specific uses are described for each of the

CC polynucleotides, based on which tissues they are most highly expressed

CC in, and include developing products for the diagnosis or treatment of

CC cancer, tumours, neurodegenerative disorders, developmental abnormalities

CC and foetal deficiencies, blood disorders, diseases of the immune system,

CC autoimmune diseases, hepatic and renal disease, inflammation,

CC allergies, Alzheimer's and behavioural disorders, schizophrenia,

CC osteoporosis, arthritis, infections, AIDS, spinal cord injuries,

CC transplant rejection, diabetes, asthma, sepsis, acne, psoriasis,

CC cardiovascular disorders, reproductive disorders, gastrointestinal

CC disorders, respiratory disorders and metabolic disorders. The proteins

CC or polynucleotides can also be used as food additives or preservatives.

CC The proteins are also useful for identifying their binding partners.

CC AAA26337 to AAA26345 and AAY91450 are sequences used in the

CC exemplification of the present invention.

XX Sequence 374 AA;

Query Match 99.1%; Score 1090; DB 21; Length 374;

Best Local Similarity 99.0%; Pred. No. 1.1e-102;

Matches 201; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RLLSGQPVCRRGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI ESEDEQKLIEX 60

Db 25 RLLSGQPVCRRGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI ESEDEQKLIEX 84

QY 61 FIENLLPSDGDGFWIGLRRREKQSNSTACQDLYAWTDGSI SQFRNMYVDEPSCGSEVCVV 120

Db 85 FIENLLPSDGDGFWIGLRRREKQSNSTXCQDLYAWTDGSI SQFRNMYVDEPSCGSEVCVV 144

QY 121 MYHQPSAPAGIGGPPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGETELTTPVLPE 180

Db 145 MYHQPSAPAGIGGPPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGETELTTPVLPE 204

QY 181 ETQEDAKKTFKESREAAALNLAY 203

Db 205 ETQEDAKKTFKESREAAALNLAY 227

RESULT 7

AAY13367

ID AAY13367 standard; Protein; 382 AA.

XX

AC AAY13367;

XX

DT 25-JUN-1999 (first entry)

XX

DE Amino acid sequence of protein PRO234.

XX

KW Secreted protein; transmembrane protein; human; enterocolitis;

KW Zollinger-Ellison syndrome; gastrointestinal ulceration;

KW congenital microvillus atrophy; skin disease; cell growth;

KW abnormal keratinocyte differentiation; psoriasis; epithelial cancer;

KW Parkinson's disease; Alzheimer's disease; ALS; neuropathy;

KW fibromodulin; dermal scarring; Usher Syndrome; Atrophia areata;

KW anti-thrombotic; wound healing; tissue repair.

XX

OS Homo sapiens.

XX

PN WO9914328-A2.

XX

PD 25-MAR-1999.

XX

PF 16-SEP-1998; 98WO-US19330.

XX

PR 25-NOV-1997; 97US-0066840.

PI 17-SEP-1997; 97US-0059113.

PR 17-SEP-1997; 97US-0059115.

PR 17-SEP-1997; 97US-0059117.

PR 17-SEP-1997; 97US-0059119.

PR 17-SEP-1997; 97US-0059121.

PR 17-SEP-1997; 97US-0059122.

PR 17-SEP-1997; 97US-0059184.

PR 18-SEP-1997; 97US-0059263.

PR 18-SEP-1997; 97US-0059266.

PR 15-OCT-1997; 97US-0062125.

PR 17-OCT-1997; 97US-0062285.

PR 17-OCT-1997; 97US-0062287.

PR 21-OCT-1997; 97US-0063486.

PR 24-OCT-1997; 97US-0062814.

PR 24-OCT-1997; 97US-0062816.

PR 24-OCT-1997; 97US-0063045.

PR 24-OCT-1997; 97US-0063120.

PR 24-OCT-1997; 97US-0063121.

PR 24-OCT-1997; 97US-0063127.

PR 24-OCT-1997; 97US-0063128.

PR 27-OCT-1997; 97US-0063329.

PR 27-OCT-1997; 97US-0063327.

PR 28-OCT-1997; 97US-0063541.

PR 28-OCT-1997; 97US-0063542.

PR 28-OCT-1997; 97US-0063544.

PR 28-OCT-1997; 97US-0063549.

PR 28-OCT-1997; 97US-0063550.

PR 28-OCT-1997; 97US-0063564.

PR 29-OCT-1997; 97US-0063435.

PR 29-OCT-1997; 97US-0063704.

DR N-PSDB; AAS45934.

XX Novel nucleic acids encoding PRO polypeptides, used to diagnose the

PT presence of tumours, such as prostate and breast tumours, in mammals and

PT to screen for modulators of the compounds -

XX

PS Claim 11; Fig 20; 774pp; English.

XX

CC Sequences AAU29024-AAU29328 represent PRO polypeptides of the invention.

CC The PRO polypeptides and their associated nucleic acids can be used to

CC detect the presence of a tumour in a mammal by comparing the level of

CC expression of a PRO polypeptide in a test sample of cells from the animal

CC and a control sample of normal cells, whereby a higher level of

CC expression in the test sample indicates the presence of a tumour in the

CC mammal. Mammals include dogs, cats, cattle, horses, sheep, pigs, goats

CC and rabbits but are preferably human. The polypeptides can be used to

CC stimulate tumour necrosis factor (TNF) alpha release from human blood,

CC when contacted with it. A specific polypeptide can be used to stimulate

CC the proliferation or differentiation of chondrocyte cells. The PRO

CC proteins can be used to determine the presence of tumours and also

CC susceptibility to tumour development, particularly adrenal, lung, colon,

CC breast, prostate, rectal, cervical, or liver tumours, in mammalian

CC subjects. The oligonucleotide probes specific for the PRO nucleic acids

CC can be used for genetic analysis of individuals with genetic disorders.

XX

SQ Sequence 382 AA;

Query Match 98.7%; Score 1086; DB 22; Length 382;

Best Local Similarity 96.2%; Pred. No. 2.8e-102;

Matches 203; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 RLLS-----GQVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIIE 52

Db 25 RLLSASDLDLRGQPVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIIE 84

QY 53 DEQKLIKFIENLLPSDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNYYVDEPS 112

Db 85 DEQKLIKFIENLLPSDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNYYVDEPS 144

QY 113 CGSEVCVMYHQPAPAGIGGPMFQWDDRCNMKNFICKYSDEKPAVPSREAEGETE 172

Db 145 CGSEVCVMYHQPAPAGIGGPMFQWDDRCNMKNFICKYSDEKPAVPSREAEGETE 204

QY 173 LTPVLPEETQEDAKKTFKESREAAALNLAY 203

Db 205 LTPVLPEETQEDAKKTFKESREAAALNLAY 235

RESULT 9

AAB80235

ID AAB80235 standard; Protein; 382 AA.

XX

AC AAB80235;

XX

DT 24-APR-2001 (first entry)

XX

DE Human PRO234 protein.

XX

KW Human; PRO; dermatological; antipsoriatic; cytostatic; antiinflammatory;

KW antiparkinsonian nootropic; neuroprotective; vulnerary; cardiant;

KW antiangiogenic; vasotropic; antiasthmatic; antirheumatic; cancer;

KW antiarthritic; antiinfertility; antidiabetic; antiviral; diabetes;

KW ophthalmological; gene therapy; skin disease; gastrointestinal disorder;

KW ischaemia; inflammation.

XX

OS Homo sapiens.

XX

PN WO200104311-A1.

XX

PD 18-JAN-2001.

XX

PF 22-FEB-2000; 2000WO-US04414.

XX

PR 07-JUL-1999; 99US-0143048.

PR 26-JUL-1999; 99US-0145698.

PR 28-JUL-1999; 99US-0146222.

PR 08-SEP-1999; 99WO-US20594.

PR 13-SEP-1999; 99WO-US20944.

PR 15-SEP-1999; 99WO-US21090.

PR 15-SEP-1999; 99WO-US21547.

PR 05-OCT-1999; 99WO-US23089.

PR 29-NOV-1999; 99WO-US28214.

PR 30-NOV-1999; 99WO-US28313.

PR 16-DEC-1999; 99WO-US30095.

PR 20-DEC-1999; 99WO-US30911.

PR 20-DEC-1999; 99WO-US30999.

PR 05-JAN-2000; 99WO-US00219.

XX (GETH) GENENTECH INC.

PA Ashkenazi AJ, Botstein D, Desnoyers L, Eaton DL, Ferrara N;

XX Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME, Goddard A;

PI Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ, Kljavin IJ;

PI Mather JP, Pan J, Paoni NF, Roy MA, Stewart TA, Tumas D;

PI Williams PM, Wood WI;

XX WPI; 2001-081051/09.

DR N-PSDB; AAF72396.

XX

PT Sixty one nucleic acids encoding PRO polypeptides which are useful in

PT the treatment of skin diseases (e.g. psoriasis), cancers (e.g. lung

PT squamous cell carcinoma) and neurodegenerative diseases (e.g.

PT Alzheimer's disease) -

XX

PS Claim 1; Fig 50; 393pp; English.

XX

CC The present sequence is one of sixty one novel secreted and

CC transmembrane PRO polypeptides. The PRO polypeptides are

CC useful for treating skin diseases (e.g. psoriasis), cancers (e.g. lung

CC squamous cell carcinoma), gastrointestinal disorders (e.g.

CC enterocolitis), neurodegenerative diseases (e.g. Alzheimer's disease,

CC Parkinson's disease), wound repair, cardiovascular disorders (e.g.

CC endometrial bleeding angiogenesis, ischaemias such as coronary

CC ischaemia, atherosclerosis), inflammatory disorders (e.g. asthma,

CC rheumatoid arthritis, multiple sclerosis), infertility, AIDS and

CC diabetes and retinal disorders such as retinitis pigmentosum.

CC The PRO nucleic acids have applications in molecular biology, including

CC use as hybridization probes, and in chromosome and gene mapping.

XX

SQ Sequence 382 AA;

Query Match 98.7%; Score 1086; DB 22; Length 382;

Best Local Similarity 96.2%; Pred. No. 2.8e-102;

Matches 203; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 RLLS-----GQVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIIE 52

Db 25 RLLSASDLDLRGQPVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIIE 84

QY 53 DEQKLIKFIENLLPSDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNYYVDEPS 112

Db 85 DEQKLIKFIENLLPSDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNYYVDEPS 144

QY 113 CGSEVCVMYHQPAPAGIGGPMFQWDDRCNMKNFICKYSDEKPAVPSREAEGETE 172

Db 145 CGSEVCVMYHQPAPAGIGGPMFQWDDRCNMKNFICKYSDEKPAVPSREAEGETE 204

QY 173 LTPVLPEETQEDAKKTFKESREAAALNLAY 203

Db 205 LTPVLPEETQEDAKKTFKESREAAALNLAY 235

RESULT 10

ABU69645

ID ABU69645 standard; Protein; 382 AA.

XX

CC protein electrophoresis purposes. The anti-PRO antibodies may be used
CC in diagnostic assays for PRO, or for the affinity purification of PRO
CC from recombinant cell culture or natural sources. This is the amino
CC acid sequence of a novel human PRO polypeptide.

SQ Sequence 382 AA;

Query Match 98.7%; Score 1086; DB 24; Length 382;
Best Local Similarity 96.2%; Pred. No. 2.8e-102;
Matches 203; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 RLLS-----GQVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
||| |
Db 25 RLLSASDLRLGGQPCVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 84
||| |

QY 53 DEQKLEKFIENLLPSDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNWWYDEPS 112
||| |
Db 85 DEQKLEKFIENLLPSDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNWWYDEPS 144
||| |

QY 113 CGSEVCVMYHQPAPAGIGGYPYFQWDDRCNMKNFICKYSDEKPAVPSREAEGETE 172
||| |
Db 145 CGSEVCVMYHQPAPAGIGGYPYFQWDDRCNMKNFICKYSDEKPAVPSREAEGETE 204
||| |

QY 173 LTPVLPETQEEDAKKTFKESREAAALNLAY 203
||| |
Db 205 LTPVLPETQEEDAKKTFKESREAAALNLAY 235
||| |

RESULT 11
ABU71121

ID ABU71121 standard; Protein; 382 AA.

XX

AC ABU71121;

XX

DT 10-JUN-2003 (first entry)

XX

DE Human PRO234 protein.

XX

KW Human; PRO; secreted; transmembrane; cytostatic; TNF-alpha; blood;
KW tumour necrosis factor alpha release; chondrocyte cell; proliferation;
KW differentiation; tumour; gene therapy.

XX

OS Homo sapiens.

XX

PN US2003036143-A1.

XX

PD 20-FEB-2003.

XX

PF 02-JUL-2002; 2002US-0187600.

XX

PR 16-SEP-1998; 98WO-US19330.
PR 07-OCT-1998; 98WO-US21141.
PR 01-DEC-1998; 98WO-US25108.
PR 08-MAR-1999; 99WO-US05028.
PR 14-MAY-1999; 99WO-US10733.
PR 02-JUN-1999; 99WO-US12252.
PR 01-SEP-1999; 99WO-US20111.
PR 15-SEP-1999; 99WO-US21090.
PR 01-DEC-1999; 99WO-US28301.
PR 02-DEC-1999; 99WO-US28551.
PR 30-DEC-1999; 99WO-US31274.
PR 05-JAN-2000; 2000WO-US00219.
PR 18-FEB-2000; 2000WO-US04341.
PR 18-FEB-2000; 2000WO-US04342.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US05004.
PR 01-MAR-2000; 2000WO-US05601.
PR 02-MAR-2000; 2000WO-US05841.
PR 15-MAR-2000; 2000WO-US06884.
PR 30-MAR-2000; 2000WO-US08439.
PR 17-MAY-2000; 2000WO-US13705.
PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.

PR 02-JUN-2000; 2000WO-US15264.
PR 28-JUL-2000; 2000WO-US20710.
PR 24-AUG-2000; 2000WO-US23328.
PR 08-NOV-2000; 2000WO-US30952.
PR 01-DEC-2000; 2000WO-US32678.
PR 20-DEC-2000; 2000WO-US34956.
PR 28-FEB-2001; 2001WO-US06520.
PR 01-JUN-2001; 2001WO-US17800.
PR 20-JUN-2001; 2001WO-US19692.
PR 29-JUN-2001; 2001WO-US21066.
PR 09-JUL-2001; 2001WO-US21735.
PR 29-AUG-2001; 2001WO-US27099.
PR 18-SEP-1997; 97US-059263P.
PR 18-SEP-1997; 97US-059266P.
PR 17-OCT-1997; 97US-062250P.
PR 21-OCT-1997; 97US-063486P.
PR 24-OCT-1997; 97US-063120P.
PR 24-OCT-1997; 97US-063121P.
PR 28-OCT-1997; 97US-063540P.
PR 28-OCT-1997; 97US-063541P.
PR 28-OCT-1997; 97US-063544P.
PR 28-OCT-1997; 97US-063564P.
PR 29-OCT-1997; 97US-063734P.
PR 31-OCT-1997; 97US-063870P.
PR 31-OCT-1997; 97US-064103P.
PR 13-NOV-1997; 97US-065311P.
PR 21-NOV-1997; 97US-066120P.
PR 24-NOV-1997; 97US-066466P.
PR 11-DEC-1997; 97US-066772P.
PR 12-DEC-1997; 97US-069425P.
PR 17-DEC-1997; 97US-069870P.
PR 18-DEC-1997; 97US-068017P.
PR 10-MAR-1998; 98US-077450P.
PR 11-MAR-1998; 98US-077632P.
PR 11-MAR-1998; 98US-077649P.
PR 20-MAR-1998; 98US-078886P.
PR 20-MAR-1998; 98US-078939P.
PR 27-MAR-1998; 98US-079664P.
PR 27-MAR-1998; 98US-079786P.
PR 31-MAR-1998; 98US-080107P.
PR 31-MAR-1998; 98US-080194P.
PR 01-APR-1998; 98US-080327P.
PR 01-APR-1998; 98US-080333P.
PR 08-APR-1998; 98US-081049P.
PR 08-APR-1998; 98US-081070P.
PR 09-APR-1998; 98US-081195P.
PR 15-APR-1998; 98US-081838P.
PR 21-APR-1998; 98US-082568P.
PR 21-APR-1998; 98US-082569P.
PR 22-APR-1998; 98US-082704P.
PR 22-APR-1998; 98US-082797P.
PR 28-APR-1998; 98US-083322P.
PR 29-APR-1998; 98US-083495P.
PR 29-APR-1998; 98US-083496P.
PR 29-APR-1998; 98US-083499P.
PR 29-APR-1998; 98US-083559P.
PR 05-MAY-1998; 98US-084366P.
PR 06-MAY-1998; 98US-084414P.
PR 07-MAY-1998; 98US-084639P.
PR 07-MAY-1998; 98US-084640P.
PR 07-MAY-1998; 98US-084643P.
PR 15-MAY-1998; 98US-085579P.
PR 15-MAY-1998; 98US-085580P.
PR 15-MAY-1998; 98US-085582P.
PR 15-MAY-1998; 98US-085700P.
PR 18-MAY-1998; 98US-086023P.
PR 22-MAY-1998; 98US-086392P.
PR 22-MAY-1998; 98US-086486P.
PR 28-MAY-1998; 98US-087098P.
PR 28-MAY-1998; 98US-087208P.
PR 02-JUN-1998; 98US-087609P.
PR 02-JUN-1998; 98US-087759P.

PR 03-JUN-1998; 98US-087827P.
PR 04-JUN-1998; 98US-088025P.
PR 04-JUN-1998; 98US-088028P.
PR 04-JUN-1998; 98US-088029P.
PR 04-JUN-1998; 98US-088033P.
PR 04-JUN-1998; 98US-088326P.
PR 05-JUN-1998; 98US-088167P.
PR 05-JUN-1998; 98US-088202P.
PR 05-JUN-1998; 98US-088212P.
PR 05-JUN-1998; 98US-088217P.
PR 09-JUN-1998; 98US-088655P.
PR 10-JUN-1998; 98US-088722P.
PR 10-JUN-1998; 98US-088738P.
PR 10-JUN-1998; 98US-088740P.
PR 10-JUN-1998; 98US-088811P.
PR 10-JUN-1998; 98US-088824P.
PR 10-JUN-1998; 98US-088825P.
PR 10-JUN-1998; 98US-088826P.
PR 11-JUN-1998; 98US-088861P.
PR 11-JUN-1998; 98US-088863P.
PR 11-JUN-1998; 98US-088876P.
PR 12-JUN-1998; 98US-089090P.
PR 12-JUN-1998; 98US-089105P.
PR 16-JUN-1998; 98US-089512P.
PR 16-JUN-1998; 98US-089514P.
PR 17-JUN-1998; 98US-089538P.
PR 17-JUN-1998; 98US-089598P.
PR 17-JUN-1998; 98US-089653P.
PR 18-JUN-1998; 98US-089908P.
PR 19-JUN-1998; 98US-089952P.
PR 22-JUN-1998; 98US-090246P.
PR 22-JUN-1998; 98US-090252P.
PR 22-JUN-1998; 98US-090254P.
PR 24-JUN-1998; 98US-090429P.
PR 24-JUN-1998; 98US-090435P.
PR 24-JUN-1998; 98US-090444P.
PR 24-JUN-1998; 98US-090461P.
PR 24-JUN-1998; 98US-090535P.
PR 24-JUN-1998; 98US-090540P.
PR 25-JUN-1998; 98US-090676P.
PR 25-JUN-1998; 98US-090678P.
PR 25-JUN-1998; 98US-090688P.
PR 25-JUN-1998; 98US-090690P.
PR 25-JUN-1998; 98US-090694P.
PR 25-JUN-1998; 98US-090695P.
PR 25-JUN-1998; 98US-090696P.
PR 26-JUN-1998; 98US-090862P.
PR 26-JUN-1998; 98US-090863P.
PR 26-JUN-1998; 98US-091010P.
PR 01-JUL-1998; 98US-091359P.
PR 01-JUL-1998; 98US-091544P.
PR 02-JUL-1998; 98US-091478P.
PR 02-JUL-1998; 98US-091486P.
PR 02-JUL-1998; 98US-091626P.
PR 02-JUL-1998; 98US-091628P.
PR 02-JUL-1998; 98US-091632P.
PR 24-JUL-1998; 98US-094006P.
PR 04-AUG-1998; 98US-095282P.
PR 10-AUG-1998; 98US-095998P.
PR 10-AUG-1998; 98US-096012P.
PR 17-AUG-1998; 98US-096757P.
PR 17-AUG-1998; 98US-096766P.
PR 17-AUG-1998; 98US-096867P.
PR 17-AUG-1998; 98US-096891P.
PR 17-AUG-1998; 98US-096897P.
PR 18-AUG-1998; 98US-096949P.
PR 18-AUG-1998; 98US-096959P.
PR 18-AUG-1998; 98US-097022P.
PR 26-AUG-1998; 98US-097952P.
PR 26-AUG-1998; 98US-097954P.
PR 26-AUG-1998; 98US-097955P.
PR 26-AUG-1998; 98US-097971P.
PR 26-AUG-1998; 98US-097974P.

PR 26-AUG-1998; 98US-098014P.
PR 01-SEP-1998; 98US-098716P.
PR 01-SEP-1998; 98US-098723P.
PR 02-SEP-1998; 98US-098803P.
PR 02-SEP-1998; 98US-098821P.
PR 02-SEP-1998; 98US-098843P.
PR 09-SEP-1998; 98US-099602P.
PR 10-SEP-1998; 98US-099741P.
PR 10-SEP-1998; 98US-099754P.
PR 10-SEP-1998; 98US-099763P.
PR 10-SEP-1998; 98US-099812P.

Query Match 98.7%; Score 1086; DB 24; Length 382;
Best Local Similarity 96.2%; Pred. No. 2.8e-102;
Matches 203; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
Qy 1 RLLS-----GQVCRGGTQPCVKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI ESE 52
Db 25 RLLSASDLRLRGQPCVKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI ESE 84
Qy 53 DEQKLI EKFIENLLPSDGD F WIGLRRRREKQSNSTACQDLYAWTDGSI SQFRN WYVDEPS 112
Db 85 DEQKLI EKFIENLLPSDGD F WIGLRRRREKQSNSTACQDLYAWTDGSI SQFRN WYVDEPS 144
Qy 113 CGSEVCVVMYHQP SAPAGIGG PYMFQWNDRCNMKNFICKYSDEKPAVPSREAEGETE 172
Db 145 CGSEVCVVMYHQP SAPAGIGG PYMFQWNDRCNMKNFICKYSDEKPAVPSREAEGETE 204
Qy 173 LTT PVLPEETQEEDAKTTFKESREAAALN LAY 203
Db 205 LTT PVLPEETQEEDAKTTFKESREAAALN LAY 235

RESULT 12

ABU71468
ID ABU71468 standard; Protein; 382 AA.
XX
AC ABU71468;
XX
DT 10-JUN-2003 (first entry)
XX
DE Human PRO polypeptide #24.
XX
KW Human; secreted and transmembrane protein; PRO polypeptide; cancer;
KW Alzheimer's disease; ischaemia; cytostatic; nootropic; vasotropic;
KW neuroprotective.
XX
OS Homo sapiens.
XX
PN US2002192659-A1.
XX
PD 19-DEC-2002.
XX
PF 10-JUL-2001; 2001US-0902853.
XX
PR 10-SEP-1998; 98WO-US18824.
PR 14-SEP-1998; 98WO-US19177.
PR 16-SEP-1998; 98WO-US19330.
PR 17-SEP-1998; 98WO-US19437.
PR 01-DEC-1998; 98WO-US25108.
PR 08-SEP-1999; 99WO-US20594.
PR 13-SEP-1999; 99WO-US20944.
PR 15-SEP-1999; 99WO-US21090.
PR 15-SEP-1999; 99WO-US21547.
PR 05-OCT-1999; 99WO-US23089.
PR 01-DEC-1999; 99WO-US28301.
PR 02-DEC-1999; 99WO-US28564.
PR 02-DEC-1999; 99WO-US28565.
PR 16-DEC-1999; 99WO-US30095.
PR 20-DEC-1999; 99WO-US30911.
PR 20-DEC-1999; 99WO-US30999.
PR 05-JAN-2000; 2000WO-US00219.
PR 11-FEB-2000; 2000WO-US03565.

```
PR 22-FEB-2000; 2000WO-US04414.
PR 28-JUL-2000; 2000WO-US20710.
PR 24-AUG-2000; 2000WO-US23328.
PR 17-SEP-1997; 97US-059113P.
PR 17-SEP-1997; 97US-059115P.
PR 17-SEP-1997; 97US-059117P.
PR 18-SEP-1997; 97US-059266P.
PR 15-OCT-1997; 97US-062125P.
PR 17-OCT-1997; 97US-062285P.
PR 17-OCT-1997; 97US-062287P.
PR 21-OCT-1997; 97US-063486P.
PR 24-OCT-1997; 97US-062814P.
PR 24-OCT-1997; 97US-062816P.
XX
PA (GETH ) GENENTECH INC.
XX
PI Ashkenazi A, Botstein D, Desnoyers L, Eaton DL, Ferrara N;
PI Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME, Goddard A;
PI Godowski PJ, Grimaldi JC, Gurney AL, Hillan KJ, Kljavin IJ;
PI Mather JP, Pan J, Paoni NF, Roy MA, Stewart TA, Tumas D;
PI Williams PM, Wood WI;
XX
DR WPI; 2003-361832/34.
DR N-PSDB; ACA58405.
XX
PT New isolated nucleic acid encoding a PRO polypeptide, e.g. PRO245 or
PT PRO1868, useful in molecular biology, chromosome and gene mapping, in
PT generating antisense RNA and DNA, and in gene therapy -
XX
PS Claim 12; Fig 50; 474pp; English.
XX
CC The present invention relates to the isolation of novel human secreted
CC and transmembrane proteins (PRO polypeptides), and the polynucleotide
CC sequences encoding them. The polynucleotide sequences are useful in
CC molecular biology, as hybridisation probes, in chromosome and gene
CC mapping, in generating antisense RNA and DNA, and in gene therapy. The
CC polynucleotide sequences may also be used in preparing PRO polypeptides
CC by recombinant techniques, and in generating either transgenic animals
CC or knock-out animals which, in turn, are useful in the development and
CC screening of therapeutically useful reagents. The PRO polypeptides or
CC their antibodies are useful in preparing a medicament for treating a
CC condition responsive to the polypeptide or antibody, such as cancer,
CC Alzheimer's disease or ischaemia, and in various diagnostic assays.
CC ABU71445-ABU71505 represent human PRO polypeptides of the invention.
XX
SQ Sequence 382 AA;
Query Match 98.7%; Score 1086; DB 24; Length 382;
Best Local Similarity 96.2%; Pred. No. 2.8e-102;
Matches 203; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
QY 1 RLLS-----GQVCRGGTQPCYKVIYFHTSRRLNFEFEAKEACRRDGGQLVSIESE 52
Db |||||
25 RLLSASDLRLRGQPVCRGGTQPCYKVIYFHTSRRLNFEFEAKEACRRDGGQLVSIESE 84
QY 53 DEQKLIKFIENLLPSDGDGFWIGLRRRREEKQSNSTACQDLYAWTDGSIQFRNMYVDEPS 112
Db |||||
85 DEQKLIKFIENLLPSDGDGFWIGLRRRREEKQSNSTACQDLYAWTDGSIQFRNMYVDEPS 144
QY 113 CGSEVCVVMYHQPSAPAGIGGPFYMFQWDDRCNMKNFNICKYSDEKPAVPSREAEGEETE 172
Db |||||
145 CGSEVCVVMYHQPSAPAGIGGPFYMFQWDDRCNMKNFNICKYSDEKPAVPSREAEGEETE 204
QY 173 LTTPLVPEETQEEDAKTKTFKESREAAALNLAY 203
Db |||||
205 LTTPLVPEETQEEDAKTKTFKESREAAALNLAY 235
RESULT 13
ABU71914
ID ABU71914 standard; Protein; 382 AA.
XX
AC ABU71914;
```

```
XX
DT 12-JUN-2003 (first entry)
XX
DE Human secreted/transmembrane protein PRO234.
XX
KW Human; secreted protein; transmembrane protein; PRO;
KW gene therapy; chromosome identification; chromosome marker.
XX
OS Homo sapiens.
XX
PN US2003003530-A1.
XX
PD 02-JAN-2003.
XX
PF 11-JUL-2001; 2001US-0904011.
XX
PR 10-SEP-1998; 98WO-US18824.
PR 14-SEP-1998; 98WO-US19177.
PR 16-SEP-1998; 98WO-US19330.
PK 17-SEP-1998; 98WO-US19437.
PR 01-DEC-1998; 98WO-US25108.
PR 08-SEP-1999; 99WO-US20594.
PR 13-SEP-1999; 99WO-US20944.
PR 15-SEP-1999; 99WO-US21090.
PR 15-SEP-1999; 99WO-US21547.
PR 05-OCT-1999; 99WO-US23089.
PR 29-NOV-1999; 99WO-US28214.
PR 30-NOV-1999; 99WO-US28313.
PR 01-DEC-1999; 99WO-US28301.
PR 02-DEC-1999; 99WO-US28564.
PR 02-DEC-1999; 99WO-US28565.
PR 16-DEC-1999; 99WO-US30095.
PR 20-DEC-1999; 99WO-US30911.
PR 20-DEC-1999; 99WO-US30999.
PR 05-JAN-2000; 2000WO-US00219.
PR 11-FEB-2000; 2000WO-US03565.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US05004.
PR 02-MAR-2000; 2000WO-US05841.
PR 20-MAR-2000; 2000WO-US07377.
PR 30-MAR-2000; 2000WO-US08439.
PR 22-MAY-2000; 2000WO-US14042.
PR 02-JUN-2000; 2000WO-US15264.
PR 28-JUL-2000; 2000WO-US20710.
PR 24-AUG-2000; 2000WO-US23328.
PR 17-SEP-1997; 97US-059113P.
PR 17-SEP-1997; 97US-059115P.
PR 17-SEP-1997; 97US-059117P.
PR 17-SEP-1997; 97US-059119P.
PR 17-SEP-1997; 97US-059121P.
PR 17-SEP-1997; 97US-059122P.
PR 17-SEP-1997; 97US-059184P.
PR 18-SEP-1997; 97US-059263P.
PR 18-SEP-1997; 97US-059266P.
PR 15-OCT-1997; 97US-062125P.
PR 17-OCT-1997; 97US-062285P.
PR 17-OCT-1997; 97US-062287P.
PR 21-OCT-1997; 97US-063486P.
PR 24-OCT-1997; 97US-062814P.
PR 24-OCT-1997; 97US-062816P.
PR 24-OCT-1997; 97US-063045P.
PR 24-OCT-1997; 97US-063120P.
PR 24-OCT-1997; 97US-063121P.
PR 24-OCT-1997; 97US-063127P.
PR 24-OCT-1997; 97US-063128P.
PR 27-OCT-1997; 97US-063327P.
PR 27-OCT-1997; 97US-063329P.
PR 28-OCT-1997; 97US-063541P.
PR 28-OCT-1997; 97US-063542P.
PR 28-OCT-1997; 97US-063544P.
PR 28-OCT-1997; 97US-063549P.
PR 28-OCT-1997; 97US-063550P.
PR 28-OCT-1997; 97US-063564P.
```

PR 29-OCT-1997; 97US-0634335P.
PR 29-OCT-1997; 97US-063704P.
PR 29-OCT-1997; 97US-063732P.
PR 29-OCT-1997; 97US-063734P.
PR 29-OCT-1997; 97US-063735P.
PR 29-OCT-1997; 97US-063738P.
PR 29-OCT-1997; 97US-064215P.
PR 31-OCT-1997; 97US-063870P.
PR 31-OCT-1997; 97US-064103P.
PR 03-NOV-1997; 97US-064248P.
PR 07-NOV-1997; 97US-064809P.
PR 12-NOV-1997; 97US-065186P.
PR 17-NOV-1997; 97US-065846P.
PR 18-NOV-1997; 97US-065693P.
PR 21-NOV-1997; 97US-066120P.
PR 21-NOV-1997; 97US-066364P.
PR 24-NOV-1997; 97US-066453P.
PR 24-NOV-1997; 97US-066466P.
PR 24-NOV-1997; 97US-066511P.
PR 24-NOV-1997; 97US-066770P.
PR 24-NOV-1997; 97US-066772P.
PR 18-SEP-2000; 2000US-0665350.
XX
PA (GETH) GENENTECH INC.
XX
PI Ashkenazi A, Botstein D, Desnoyers L, Eaton DL, Ferrara N;
PI Filvaroff E, Fong S, Gerber H, Gerritsen ME, Goddard A;
PI Godowski PJ, Grimaldi JC, Gurney AL, Hillan KJ, Kljavin IJ;
PI Mather JP, Pan J, Paoni NF, Roy MA, Stewart TA, Tumas D;
PI Williams PM, Wood WI;
XX
DR WPI; 2003-329602/31.
DR N-PSDB; ACA60112.
XX
XX New transmembrane polypeptides and nucleic acids encoding the
PT polypeptides, useful in gene therapy, in chromosome identification, as
PT chromosome markers, in generating probes and in tissue typing
XX
PS Claim 12; Fig 50; 484pp; English.
XX
CC The invention relates to an isolated nucleic acid with at least 80%
CC nucleic acid sequence identity to a nucleotide sequence encoding one of
CC 61 secreted/transmembrane polypeptides, or PRO polypeptides or encoding a
CC PRO protein extracellular domain. Also included are a vector comprising
CC the PRO nucleic acid, a host cell comprising the vector, producing a PRO
CC polypeptide (by culturing the host cell for the expression of the PRO
CC polypeptide, and recovering the PRO polypeptide from the cell culture),
CC an isolated PRO polypeptide (having at least 80% sequence identity
CC to: (a) an amino acid sequence selected from the 61 PRO proteins;
CC (b) an amino acid sequence encoded by a nucleic acid molecule deposited
CC with an ATCC number (detailed in the specification); or (c) an
CC extracellular domain of a PRO polypeptide or to a PRO polypeptide lacking
CC its associated signal peptide), a chimaeric molecule comprising a PRO
CC polypeptide of fused to a heterologous amino acid sequence, an anti-PRO
CC antibody, detecting a PRO245 or PRO1868 in a sample suspected of
CC containing the polypeptide, linking a bioactive molecule to a cell
CC expressing a PRO245 or PRO1868 and modulating at least one biological
CC activity of a cell expressing a PRO245 or PRO1868. Nucleic acids which
CC encode PRO can be used to generate either transgenic animals or knock-out
CC animals which may be used in the development and screening of
CC therapeutically useful reagents. The nucleic acids may also be used in
CC gene therapy, in chromosome identification, as chromosome markers, or in
CC generating probes. The PRO polypeptides are useful as molecular markers
CC for protein electrophoresis, and the isolated nucleic acids may be used
CC for recombinantly expressing those markers. The PRO polypeptides and
CC nucleic acids may also be used in tissue typing. Anti-PRO antibodies
CC are useful in diagnostic assays for PRO, and in affinity purification
CC of PRO from recombinant cell culture or natural sources. The
XX present sequence represents a PRO protein.

SQ Sequence 382 AA;

Query Match

98.7%; Score 1086; DB 24; Length 382;

Best Local Similarity 96.2%; Pred. No. 2.8e-102;
Matches 203; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 RLLS-----GQPVCRGGTQPCYKVIYFHDTSRRLNFEAEAKRRDGGQLVSI ESE 52
|||
Db 25 RLLSASDLRLRGQPVCRGGTQPCYKVIYFHDTSRRLNFEAEAKRRDGGQLVSI ESE 84
|||
QY 53 DEQKLI EKFIENLLPSDGFWIGLRRREKQSNSTACQDLYAWTDGSI SQFRNMYVDEPS 112
|||||
Db 85 DEQKLI EKFIENLLPSDGFWIGLRRREKQSNSTACQDLYAWTDGSI SQFRNMYVDEPS 144
|||||
QY 113 CGSEVCVVMYHQPSAPAGIGGPPYMFQWNDRCNMKNFICKYSDEKPAVPSREAEGETE 172
|||||
Db 145 CGSEVCVVMYHQPSAPAGIGGPPYMFQWNDRCNMKNFICKYSDEKPAVPSREAEGETE 204
|||||
QY 173 LTTPLPEETQEEDAKKTFKESREAAALNLAY 203
|||||
Db 205 LTTPLPEETQEEDAKKTFKESREAAALNLAY 235

RESULT 14
ABU65578
ID ABU65578 standard; Protein; 382 AA.
XX
AC ABU65578;
XX
DT 19-MAY-2003 (first entry)
XX
DE Human secreted/transmembrane protein, SEQ ID 20.
XX
KW Human; PRO; secreted protein; transmembrane protein;
KW cytostatic; antiarthritic; osteopathic; adrenal tumour; lung tumour;
KW colon tumour; breast tumour; prostate tumour; rectal tumour;
KW cervical tumour; liver tumour; TNF-alpha release; arthritis;
KW tumour necrosis factor alpha; chondrocyte cell; bone disorder;
KW cartilage disorder; sports injury.
XX
OS Homo sapiens.
XX
PN US2003036156-A1.
XX
PD 20-FEB-2003.
XX
PF 02-JUL-2002; 2002US-0188767.
XX
PR 16-SEP-1998; 98WO-US19330.
PR 07-OCT-1998; 98WO-US21141.
PR 01-DEC-1998; 98WO-US25108.
PR 08-MAR-1999; 99WO-US05028.
PR 14-MAY-1999; 99WO-US10733.
PR 02-JUN-1999; 99WO-US12252.
PR 01-SEP-1999; 99WO-US20111.
PR 15-SEP-1999; 99WO-US21090.
PR 01-DEC-1999; 99WO-US28301.
PR 02-DEC-1999; 99WO-US28551.
PR 30-DEC-1999; 99WO-US31274.
PR 05-JAN-2000; 2000WO-US00219.
PR 18-FEB-2000; 2000WO-US04341.
PR 18-FEB-2000; 2000WO-US04342.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US05004.
PR 01-MAR-2000; 2000WO-US05601.
PR 15-MAR-2000; 2000WO-US05841.
PR 30-MAR-2000; 2000WO-US06884.
PR 17-MAY-2000; 2000WO-US08439.
PR 22-MAY-2000; 2000WO-US13705.
PR 30-MAY-2000; 2000WO-US14042.
PR 02-JUN-2000; 2000WO-US14941.
PR 28-JUL-2000; 2000WO-US15264.
PR 24-AUG-2000; 2000WO-US20710.
PR 08-NOV-2000; 2000WO-US23328.
PR 01-DEC-2000; 2000WO-US30952.
PR 01-DEC-2000; 2000WO-US32678.

PR 20-DEC-2000; 2000WO-US34956.
PR 28-FEB-2001; 2001WO-US06520.
PR 01-JUN-2001; 2001WO-US17800.
PR 20-JUN-2001; 2001WO-US19692.
PR 29-JUN-2001; 2001WO-US21066.
PR 09-JUL-2001; 2001WO-US21735.
PR 29-AUG-2001; 2001WO-US27099.
PR 18-SEP-1997; 97US-059263P.
PR 18-SEP-1997; 97US-059266P.
PR 17-OCT-1997; 97US-062250P.
PR 21-OCT-1997; 97US-063486P.
PR 24-OCT-1997; 97US-063120P.
PR 24-OCT-1997; 97US-063121P.
PR 28-OCT-1997; 97US-063540P.
PR 28-OCT-1997; 97US-063541P.
PR 28-OCT-1997; 97US-063544P.
PR 28-OCT-1997; 97US-063564P.
PR 29-OCT-1997; 97US-063734P.
PR 31-OCT-1997; 97US-063870P.
PR 31-OCT-1997; 97US-064103P.
PR 13-NOV-1997; 97US-065311P.
PR 21-NOV-1997; 97US-066120P.
PR 24-NOV-1997; 97US-066466P.
PR 24-NOV-1997; 97US-066772P.
PR 11-DEC-1997; 97US-069335P.
PR 12-DEC-1997; 97US-069425P.
PR 17-DEC-1997; 97US-069870P.
PR 18-DEC-1997; 97US-068017P.
PR 10-MAR-1998; 98US-077450P.
PR 11-MAR-1998; 98US-077632P.
PR 11-MAR-1998; 98US-077649P.
PR 20-MAR-1998; 98US-078886P.
PR 20-MAR-1998; 98US-078939P.
PR 27-MAR-1998; 98US-079664P.
PR 27-MAR-1998; 98US-079786P.
PR 31-MAR-1998; 98US-080107P.
PR 31-MAR-1998; 98US-080194P.
PR 01-APR-1998; 98US-080327P.
PR 01-APR-1998; 98US-080333P.
PR 08-APR-1998; 98US-081049P.
PR 08-APR-1998; 98US-081070P.
PR 09-APR-1998; 98US-081195P.
PR 15-APR-1998; 98US-081838P.
PR 21-APR-1998; 98US-082568P.
PR 21-APR-1998; 98US-082569P.
PR 22-APR-1998; 98US-082704P.
PR 22-APR-1998; 98US-082797P.
PR 28-APR-1998; 98US-083322P.
PR 29-APR-1998; 98US-083495P.
PR 29-APR-1998; 98US-083496P.
PR 29-APR-1998; 98US-083499P.
PR 29-APR-1998; 98US-083559P.
PR 05-MAY-1998; 98US-084366P.
PR 06-MAY-1998; 98US-084414P.
PR 07-MAY-1998; 98US-084639P.
PR 07-MAY-1998; 98US-084640P.
PR 07-MAY-1998; 98US-084643P.
PR 15-MAY-1998; 98US-085579P.
PR 15-MAY-1998; 98US-085580P.
PR 15-MAY-1998; 98US-085582P.
PR 15-MAY-1998; 98US-085700P.
PR 18-MAY-1998; 98US-086023P.
PR 22-MAY-1998; 98US-086392P.
PR 22-MAY-1998; 98US-086486P.
PR 28-MAY-1998; 98US-087098P.
PR 28-MAY-1998; 98US-087208P.
PR 02-JUN-1998; 98US-087609P.
PR 02-JUN-1998; 98US-087759P.
PR 03-JUN-1998; 98US-087827P.
PR 04-JUN-1998; 98US-088025P.
PR 04-JUN-1998; 98US-088028P.
PR 04-JUN-1998; 98US-088029P.
PR 04-JUN-1998; 98US-088033P.

PR 04-JUN-1998; 98US-088326P.
PR 05-JUN-1998; 98US-088167P.
PR 05-JUN-1998; 98US-088202P.
PR 05-JUN-1998; 98US-088212P.
PR 05-JUN-1998; 98US-088217P.
PR 09-JUN-1998; 98US-088655P.
PR 10-JUN-1998; 98US-088722P.
PR 10-JUN-1998; 98US-088738P.
PR 10-JUN-1998; 98US-088740P.
PR 10-JUN-1998; 98US-088811P.
PR 10-JUN-1998; 98US-088824P.
PR 10-JUN-1998; 98US-088825P.
PR 10-JUN-1998; 98US-088826P.
PR 11-JUN-1998; 98US-088861P.
PR 11-JUN-1998; 98US-088863P.
PR 11-JUN-1998; 98US-088876P.
PR 12-JUN-1998; 98US-089090P.
PR 12-JUN-1998; 98US-089105P.
PR 16-JUN-1998; 98US-089512P.
PR 16-JUN-1998; 98US-089514P.
PR 17-JUN-1998; 98US-089538P.
PR 17-JUN-1998; 98US-089598P.
PR 17-JUN-1998; 98US-089653P.
PR 18-JUN-1998; 98US-089908P.
PR 19-JUN-1998; 98US-089952P.
PR 22-JUN-1998; 98US-090246P.
PR 22-JUN-1998; 98US-090252P.
PR 22-JUN-1998; 98US-090254P.
PR 24-JUN-1998; 98US-090429P.
PR 24-JUN-1998; 98US-090435P.
PR 24-JUN-1998; 98US-090444P.
PR 24-JUN-1998; 98US-090461P.
PR 24-JUN-1998; 98US-090535P.
PR 24-JUN-1998; 98US-090540P.
PR 25-JUN-1998; 98US-090676P.
PR 25-JUN-1998; 98US-090678P.
PR 25-JUN-1998; 98US-090688P.
PR 25-JUN-1998; 98US-090690P.
PR 25-JUN-1998; 98US-090694P.
PR 25-JUN-1998; 98US-090695P.
PR 25-JUN-1998; 98US-090696P.
PR 26-JUN-1998; 98US-090862P.
PR 26-JUN-1998; 98US-090863P.
PR 26-JUN-1998; 98US-091010P.
PR 01-JUL-1998; 98US-091359P.
PR 01-JUL-1998; 98US-091544P.
PR 02-JUL-1998; 98US-091478P.
PR 02-JUL-1998; 98US-091486P.
PR 02-JUL-1998; 98US-091626P.
PR 02-JUL-1998; 98US-091628P.
PR 02-JUL-1998; 98US-091632P.
PR 24-JUL-1998; 98US-094006P.
PR 04-AUG-1998; 98US-095282P.
PR 10-AUG-1998; 98US-095998P.
PR 10-AUG-1998; 98US-096012P.
PR 17-AUG-1998; 98US-096757P.
PR 17-AUG-1998; 98US-096766P.
PR 17-AUG-1998; 98US-096867P.
PR 17-AUG-1998; 98US-096891P.
PR 17-AUG-1998; 98US-096897P.
PR 18-AUG-1998; 98US-096949P.
PR 18-AUG-1998; 98US-096959P.
PR 18-AUG-1998; 98US-097022P.
PR 26-AUG-1998; 98US-097952P.
PR 26-AUG-1998; 98US-097954P.
PR 26-AUG-1998; 98US-097955P.
PR 26-AUG-1998; 98US-097971P.
PR 26-AUG-1998; 98US-097974P.
PR 26-AUG-1998; 98US-098014P.
PR 01-SEP-1998; 98US-098716P.
PR 01-SEP-1998; 98US-098723P.
PR 02-SEP-1998; 98US-098803P.
PR 02-SEP-1998; 98US-098821P.

PR 02-SEP-1998; 98US-098843P.
PR 09-SEP-1998; 98US-099602P.
PR 10-SEP-1998; 98US-099741P.

Query Match 98.7%; Score 1086; DB 24; Length 382;
Best Local Similarity 96.2%; Pred. No. 2.8e-102;
Matches 203; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

Qy 1 RLLS-----GQVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
Db 25 RLLSASDLDRGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 84

Qy 53 DEQKLIKFIENLLPSDGFWIGLRRREEKQNSTACQDLYAWTDGSIQFRNWWYVDEPS 112
Db 85 DEQKLIKFIENLLPSDGFWIGLRRREEKQNSTACQDLYAWTDGSIQFRNWWYVDEPS 144

Qy 113 CGSEVCVMYHQSAPAGIGGPFYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGETE 172
Db 145 CGSEVCVMYHQSAPAGIGGPFYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGETE 204

Qy 173 LTPVLPEETQEEDAKKTFKESREAAALNLAY 203
Db 205 LTPVLPEETQEEDAKKTFKESREAAALNLAY 235

RESULT 15
ABU65911
ID ABU65911 standard; Protein; 382 AA.
XX
AC ABU65911;
XX
DT 20-MAY-2003 (first entry)
XX
DE Novel human secreted and transmembrane protein PRO234.
XX
KW Human; secreted protein; transmembrane protein; cytostatic;
KW gene Therapy; TNF-Agonist-Alpha; chondrocyte stimulator; tumour;
KW adrenal tumour; lung tumour; colon tumour; breast tumour;
KW prostate tumour; rectal tumour; cervical tumour; liver tumour.
XX
OS Homo sapiens.
XX
PN US2003036157-A1.
XX
PD 20-FEB-2003.
XX
PF 02-JUL-2002; 2002US-0188769.
XX
PR 16-SEP-1998; 98WO-US19330.
PR 07-OCT-1998; 98WO-US21141.
PR 01-DEC-1998; 98WO-US25108.
PR 08-MAR-1999; 99WO-US05028.
PR 14-MAY-1999; 99WO-US10733.
PR 02-JUN-1999; 99WO-US12252.
PR 01-SEP-1999; 99WO-US20111.
PR 15-SEP-1999; 99WO-US21090.
PR 01-DEC-1999; 99WO-US28301.
PR 02-DEC-1999; 99WO-US28551.
PR 30-DEC-1999; 99WO-US31274.
PR 05-JAN-2000; 2000WO-US00219.
PR 18-FEB-2000; 2000WO-US04341.
PR 18-FEB-2000; 2000WO-US04342.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US05004.
PR 01-MAR-2000; 2000WO-US05601.
PR 02-MAR-2000; 2000WO-US05841.
PR 15-MAR-2000; 2000WO-US06884.
PR 30-MAR-2000; 2000WO-US08439.
PR 17-MAY-2000; 2000WO-US13705.
PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.
PR 02-JUN-2000; 2000WO-US15264.
PR 28-JUL-2000; 2000WO-US20710.

PR 24-AUG-2000; 2000WO-US23328.
PR 08-NOV-2000; 2000WO-US30952.
PR 01-DEC-2000; 2000WO-US32678.
PR 20-DEC-2000; 2000WO-US34956.
PR 28-FEB-2001; 2001WO-US06520.
PR 01-JUN-2001; 2001WO-US17800.
PR 20-JUN-2001; 2001WO-US19692.
PR 29-JUN-2001; 2001WO-US21066.
PR 09-JUL-2001; 2001WO-US21735.
PR 29-AUG-2001; 2001WO-US27099.
PR 18-SEP-1997; 97US-059263P.
PR 18-SEP-1997; 97US-059266P.
PR 17-OCT-1997; 97US-062250P.
PR 21-OCT-1997; 97US-063486P.
PR 24-OCT-1997; 97US-063120P.
PR 24-OCT-1997; 97US-063121P.
PR 28-OCT-1997; 97US-063540P.
PR 28-OCT-1997; 97US-063541P.
PR 28-OCT-1997; 97US-063544P.
PR 28-OCT-1997; 97US-063564P.
PR 29-OCT-1997; 97US-063734P.
PR 31-OCT-1997; 97US-063870P.
PR 31-OCT-1997; 97US-064103P.
PR 13-NOV-1997; 97US-065311P.
PR 21-NOV-1997; 97US-066120P.
PR 24-NOV-1997; 97US-066466P.
PR 24-NOV-1997; 97US-066772P.
PR 11-DEC-1997; 97US-069335P.
PR 12-DEC-1997; 97US-069425P.
PR 17-DEC-1997; 97US-069870P.
PR 18-DEC-1997; 97US-068017P.
PR 10-MAR-1998; 98US-077450P.
PR 11-MAR-1998; 98US-077632P.
PR 11-MAR-1998; 98US-077649P.
PR 20-MAR-1998; 98US-078886P.
PR 20-MAR-1998; 98US-078939P.
PR 27-MAR-1998; 98US-079664P.
PR 27-MAR-1998; 98US-079786P.
PR 31-MAR-1998; 98US-080107P.
PR 31-MAR-1998; 98US-080194P.
PR 01-APR-1998; 98US-080327P.
PR 01-APR-1998; 98US-080333P.
PR 08-APR-1998; 98US-081049P.
PR 08-APR-1998; 98US-081070P.
PR 09-APR-1998; 98US-081195P.
PR 15-APR-1998; 98US-081838P.
PR 21-APR-1998; 98US-082568P.
PR 21-APR-1998; 98US-082569P.
PR 22-APR-1998; 98US-082704P.
PR 22-APR-1998; 98US-082797P.
PR 28-APR-1998; 98US-083322P.
PR 29-APR-1998; 98US-083495P.
PR 29-APR-1998; 98US-083496P.
PR 29-APR-1998; 98US-083499P.
PR 29-APR-1998; 98US-083559P.
PR 05-MAY-1998; 98US-084366P.
PR 06-MAY-1998; 98US-084414P.
PR 07-MAY-1998; 98US-084639P.
PR 07-MAY-1998; 98US-084640P.
PR 07-MAY-1998; 98US-084643P.
PR 15-MAY-1998; 98US-085579P.
PR 15-MAY-1998; 98US-085580P.
PR 15-MAY-1998; 98US-085582P.
PR 15-MAY-1998; 98US-085700P.
PR 18-MAY-1998; 98US-086023P.
PR 22-MAY-1998; 98US-086392P.
PR 22-MAY-1998; 98US-086486P.
PR 28-MAY-1998; 98US-087098P.
PR 28-MAY-1998; 98US-087208P.
PR 02-JUN-1998; 98US-087609P.
PR 02-JUN-1998; 98US-087759P.
PR 03-JUN-1998; 98US-087827P.
PR 04-JUN-1998; 98US-088025P.

PR	04-JUN-1998;	98US-088028P
PR	04-JUN-1998;	98US-088029P
PR	04-JUN-1998;	98US-088033P
PR	04-JUN-1998;	98US-0880326P
PR	05-JUN-1998;	98US-088167P
PR	05-JUN-1998;	98US-088202P
PR	05-JUN-1998;	98US-088212P
PR	05-JUN-1998;	98US-088217P
PR	09-JUN-1998;	98US-088655P
PR	10-JUN-1998;	98US-088722P
PR	10-JUN-1998;	98US-088738P
PR	10-JUN-1998;	98US-088740P
PR	10-JUN-1998;	98US-088811P
PR	10-JUN-1998;	98US-088824P
PR	10-JUN-1998;	98US-088825P
PR	10-JUN-1998;	98US-088826P
PR	11-JUN-1998;	98US-088861P
PR	11-JUN-1998;	98US-088863P
PR	11-JUN-1998;	98US-088876P
PR	12-JUN-1998;	98US-089090P
PR	12-JUN-1998;	98US-089105P
PR	16-JUN-1998;	98US-089512P
PR	16-JUN-1998;	98US-089514P
PR	17-JUN-1998;	98US-089538P
PR	17-JUN-1998;	98US-089598P
PR	17-JUN-1998;	98US-089653P
PR	18-JUN-1998;	98US-089908P
PR	19-JUN-1998;	98US-089952P
PR	22-JUN-1998;	98US-090246P
PR	22-JUN-1998;	98US-090252P
PR	22-JUN-1998;	98US-090254P
PR	24-JUN-1998;	98US-090429P
PR	24-JUN-1998;	98US-090435P
PR	24-JUN-1998;	98US-090444P
PR	24-JUN-1998;	98US-090461P
PR	24-JUN-1998;	98US-090535P
PR	24-JUN-1998;	98US-090540P
PR	25-JUN-1998;	98US-090676P
PR	25-JUN-1998;	98US-090678P
PR	25-JUN-1998;	98US-090688P
PR	25-JUN-1998;	98US-090690P
PR	25-JUN-1998;	98US-090694P
PR	25-JUN-1998;	98US-090695P
PR	25-JUN-1998;	98US-090696P
PR	26-JUN-1998;	98US-090862P
PR	26-JUN-1998;	98US-090863P
PR	26-JUN-1998;	98US-091010P
PR	01-JUL-1998;	98US-091359P
PR	01-JUL-1998;	98US-091544P
PR	02-JUL-1998;	98US-091478P
PR	02-JUL-1998;	98US-091486P
PR	02-JUL-1998;	98US-091626P
PR	02-JUL-1998;	98US-091628P
PR	02-JUL-1998;	98US-091632P
PR	24-JUL-1998;	98US-094006P
PR	04-AUG-1998;	98US-095282P
PR	10-AUG-1998;	98US-095998P
PR	10-AUG-1998;	98US-096012P
PR	17-AUG-1998;	98US-096757P
PR	17-AUG-1998;	98US-096766P
PR	17-AUG-1998;	98US-096867P
PR	17-AUG-1998;	98US-096891P
PR	17-AUG-1998;	98US-096897P
PR	18-AUG-1998;	98US-096949P
PR	18-AUG-1998;	98US-096959P
PR	18-AUG-1998;	98US-097022P
PR	26-AUG-1998;	98US-097954P
PR	26-AUG-1998;	98US-097955P
PR	26-AUG-1998;	98US-097971P
PR	26-AUG-1998;	98US-097974P
PR	26-AUG-1998;	98US-098014P
PR	01-SEP-1998;	98US-098716P

PR	01-SEP-1998;	98US-098723P.
PR	02-SEP-1998;	98US-098803P.
PR	02-SEP-1998;	98US-098821P.
PR	02-SEP-1998;	98US-098843P.
PR	09-SEP-1998;	98US-098602P.
PR	10-SEP-1998;	98US-099741P.
PR	10-SEP-1998;	98US-099754P.
PR	10-SEP-1998;	98US-099763P.

Query Match 98.7%; Score 1086; DB 24; Length 382;
Best Local Similarity 96.2%; Pred. No. 2.8e-102;
Matches 203; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

Qy	1	RLLS-----GQVCRGGTQRPCYKVIYFHTDTSRRLNFEFEAKERRDGGQLVSI	52
Db	25	RLLSASDLDRGGQPVCRGGTQRPCYKVIYFHTDTSRRLNFEFEAKERRDGGQLVSI	84
Qy	53	DEQKLEKFIENLLPSDGFWGLRRREEKQSNSTACQDLVATWDGSI	112
Db	85	DEQKLEKFIENLLPSDGFWGLRRREEKQSNSTACQDLVATWDGSI	144
Qy	113	CGSEVCVMYHQSPAPAGIGGYPMFQWDDRCNMKNPFICKYSDEKPAVPSRAE	172
Db	145	CGSEVCVMYHQSPAPAGIGGYPMFQWDDRCNMKNPFICKYSDEKPAVPSRAE	204
Qy	173	LTTPVLPEETQEEADAKTTFKESREAAALNLAY	203
Db	205	LTTPVLPEETQEEADAKTTFKESREAAALNLAY	235

Search completed: December 22, 2003, 16:10:09
Job time : 29.7569 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 22, 2003, 16:04:13 ; Search time 29.1819 Seconds
(without alignments)
1120.478 Million cell updates/sec

Title: US-09-887-855-5

Perfect score:

Sequence: 1 ATGRLLSGQPVCRGGTQRPC.....EEDAKKTFKESREAAALNLAY 206

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_19Jun03: *

1: /SIDS17gcgdata/geneseq/geneseqp-emb1/AA1980.DAT: *

2: /SIDS17gcgdata/geneseq/geneseqp-emb1/AA1981.DAT: *

3: /SIDS17gcgdata/geneseq/geneseqp-emb1/AA1982.DAT: *

4: /SIDS17gcgdata/geneseq/geneseqp-emb1/AA1983.DAT: *

5: /SIDS17gcgdata/geneseq/geneseqp-emb1/AA1984.DAT: *

6: /SIDS17gcgdata/geneseq/geneseqp-emb1/AA1985.DAT: *

7: /SIDS17gcgdata/geneseq/geneseqp-emb1/AA1986.DAT: *

8: /SIDS17gcgdata/geneseq/geneseqp-emb1/AA1987.DAT: *

9: /SIDS17gcgdata/geneseq/geneseqp-emb1/AA1988.DAT: *

10: /SIDS17gcgdata/geneseq/geneseqp-emb1/AA1989.DAT: *

11: /SIDS17gcgdata/geneseq/geneseqp-emb1/AA1990.DAT: *

12: /SIDS17gcgdata/geneseq/geneseqp-emb1/AA1991.DAT: *

13: /SIDS17gcgdata/geneseq/geneseqp-emb1/AA1992.DAT: *

14: /SIDS17gcgdata/geneseq/geneseqp-emb1/AA1993.DAT: *

15: /SIDS17gcgdata/geneseq/geneseqp-emb1/AA1994.DAT: *

16: /SIDS17gcgdata/geneseq/geneseqp-emb1/AA1995.DAT: *

17: /SIDS17gcgdata/geneseq/geneseqp-emb1/AA1996.DAT: *

18: /SIDS17gcgdata/geneseq/geneseqp-emb1/AA1997.DAT: *

19: /SIDS17gcgdata/geneseq/geneseqp-emb1/AA1998.DAT: *

20: /SIDS17gcgdata/geneseq/geneseqp-emb1/AA1999.DAT: *

21: /SIDS17gcgdata/geneseq/geneseqp-emb1/AA2000.DAT: *

22: /SIDS17gcgdata/geneseq/geneseqp-emb1/AA2001.DAT: *

23: /SIDS17gcgdata/geneseq/geneseqp-emb1/AA2002.DAT: *

24: /SIDS17gcgdata/geneseq/geneseqp-emb1/AA2003.DAT: *

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1115	100.0	374	21	AA923948	Amino acid sequenc
2	1115	100.0	374	22	AAE03651	Human extracellular
3	1115	100.0	374	23	ABG66680	Human novel polype
4	1115	100.0	374	23	ABB90203	Human polypeptide
5	1115	100.0	387	22	AAM25796	Human protein sequ
6	1105	99.1	374	21	AA91490	Human secreted pro
7	1101	98.7	382	20	AA913367	Amino acid sequenc
8	1101	98.7	382	22	AAU29033	Human PRO polypept
9	1101	98.7	382	22	AAB80235	Human PRO234 prote

ALIGNMENTS

RES.:T 1

AAV 3948

ID	AAV93948	standard:	Protein:	374 AA

AC AAY93948:

DT 03-OCT-2000 (first entry)

DE: Amino acid sequence of a lectin 883939 polypeptide

Human: lectin ss3939: chromosome 11: gene therapy

XX Homo sapiens

EH	Key	Location/Qualifiers
22		

ET	Key	Location/Qualifiers
	Pentide	1 21

```

FT      1.121  

       /note= "signal peptide"  


```

FT	Domain	22.227
FT	Domain	22.227

FT /note= "extracellular coding region"

FT Domain 228..248

FT /note= "predicted transmembrane domain"

FT	Domain	249..374
----	--------	----------

FT /note= "predicted cytoplasmic or intracellular domain"

PN WO200039296-A1.

PD 06-JUL-2000.

22-DEC-1999: 99WO-US30523

PR 23-DEC-1998: 98US-0113820

QY 1 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIIESEDEQKL 60
|||
Db 22 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIIESEDEQKL 81
QY 61 IEKFIENLLPSDGDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNMYVDEPSCGSEV 120
|||
Db 82 IEKFIENLLPSDGDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNMYVDEPSCGSEV 141
QY 121 CVMYHQPSAPAGIGGPPYMFQWNDRCNMKNFICKYSDEKPAVPSRAEAGEETELTPV 180
|||
Db 142 CVMYHQPSAPAGIGGPPYMFQWNDRCNMKNFICKYSDEKPAVPSRAEAGEETELTPV 201
QY 181 LPEETQEEDAKKTFKESREAAALNLAY 206
|||
Db 202 LPEETQEEDAKKTFKESREAAALNLAY 227

RESULT 3
ABG66680
ID ABG66680 standard; Protein; 374 AA.
XX
AC ABG66680;
XX
DT 30-AUG-2002 (first entry)
XX
DE Human novel polypeptide #15.
XX
KW Human; inflammatory condition; shock; sepsis; immune response;
KW cancer; wound healing; central nervous system disease; haematopoiesis;
KW peripheral nervous system disease; amyotrophic lateral sclerosis; tendon;
KW myeloid cell disorder; lymphoid cell disorder; platelet disorder; bone;
KW cartilage; ligament; nerve tissue; ulcer; osteoporosis; osteoarthritis;
KW bone degenerative disorder; periodontal disease; reperfusion injury;
KW lung fibrosis; liver fibrosis; autoimmune disorder; bacterial infection;
KW allergic condition; thrombolysis; thrombosis; coagulation disorder;
KW fungal infection.

OS Homo sapiens.
XX
XX WO200244340-A2.
XX
XX 06-JUN-2002.
XX
XX 30-NOV-2001; 2001WO-US47004.
XX
XX 30-NOV-2000; 2000US-0028952.
XX
XX (HYSE-) HYSEQ INC.
PI Tang YT, Goodrich RW, Liu C, Zhou P, Asundi V, Wang J, Wang D;
PI Yamazaki V, Ujwal ML, Drmanac RT;
XX
XX WPI; 2002-508509/54.
DR N-PSDB; ABK94904.
XX
XX Novel nucleic acids and polypeptides for diagnosis, treatment of
PT inflammatory, autoimmune, nervous system, myeloid or lymphoid cell
PT disorders, cancer and promoting wound healing -
XX
PS Claim 10; Page 579-580; 672pp; English.

XX The invention relates to human novel polynucleotides and associated
CC polypeptides. The polynucleotides and polypeptides are useful for
CC treating inflammatory conditions such as arthritis, nephritis, Crohn's
CC disease, ischaemia-reperfusion injury, shock, sepsis, immune responses
CC and cancer and for promoting wound healing. The sequences are used to
CC induce the proliferation of neural cells and regeneration of nerve and
CC brain tissue, and are useful for the treatment of central and peripheral
CC nervous system diseases and neuropathies, such as Alzheimer's disease,
CC parkinson's disease, Huntington's disease and amyotrophic lateral
CC sclerosis. The sequences are involved in chemotactic or chemokinetic
CC activity, regulation of haematopoiesis, treatment of myeloid or lymphoid

CC cell disorders and platelet disorders such as thrombocytopenia,
CC regeneration of bone, cartilage, tendon, ligament and/or nerve tissue
CC growth, tissue repair, healing of burns, incisions, ulcers, treatment of
CC osteoporosis, osteoarthritis, bone degenerative disorders and periodontal
CC disease. The sequences of the invention are also useful for gut
CC protection or regeneration and treatment of lung or liver fibrosis,
CC reperfusion injury in various tissues, immune deficiencies and disorders
CC including severe combined immunodeficiency (SCID), bacterial or fungal
CC infections, autoimmune disorders e.g. multiple sclerosis and myasthenia
CC gravis, allergic conditions such as asthma, thrombolysis or thrombosis
CC and coagulation disorders. Sequences ABG66666-ABG66758 represent human
CC novel polypeptides of the invention.

SQ Sequence 374 AA;

Query Match 100.0%; Score 1115; DB 23; Length 374;
Best Local Similarity 100.0%; Pred. No. 3.5e-105;
Matches 206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIIESEDEQKL 60
|||
Db 22 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIIESEDEQKL 81
QY 61 IEKFIENLLPSDGDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNMYVDEPSCGSEV 120
|||
Db 82 IEKFIENLLPSDGDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNMYVDEPSCGSEV 141
QY 121 CVMYHQPSAPAGIGGPPYMFQWNDRCNMKNFICKYSDEKPAVPSRAEAGEETELTPV 180
|||
Db 142 CVMYHQPSAPAGIGGPPYMFQWNDRCNMKNFICKYSDEKPAVPSRAEAGEETELTPV 201
QY 181 LPEETQEEDAKKTFKESREAAALNLAY 206
|||
Db 202 LPEETQEEDAKKTFKESREAAALNLAY 227

RESULT 4
ABB90203
ID ABB90203 standard; Protein; 374 AA.
XX
XX ABB90203;
XX
DT 24-MAY-2002 (first entry)
XX
DE Human polypeptide SEQ ID NO 2579.
XX
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
KW vulnery; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein.
XX
OS Homo sapiens.
XX
XX WO200190304-A2.
XX
PD 29-NOV-2001.
XX
PF 18-MAY-2001; 2001WO-US16450.
XX
XX 19-MAY-2000; 2000US-205515P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA Birse CE, Rosen CA;
PI WPI; 2002-122018/16.
XX N-PSDB; ABL90612.
PT Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
PT prevention of neural, immune system, muscular, reproductive,
PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
PT disorders -

XX PS Claim 11; SEQ ID NO 2579; 2081pp + Sequence Listing; English.

XX CC The invention relates to novel genes (ABL89449-ABL90853) and proteins

CC (ABB89040-ABB90444) useful for preventing, treating or ameliorating

CC medical conditions e.g. by protein or gene therapy. The genes are

CC isolated from a range of human tissues disclosed in the specification.

CC The nucleic acids, proteins, antibodies and (ant)agonists are useful

CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast

CC and ovarian cancer and other cancers of the adrenal gland, bone, bone

CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;

CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune

CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's

CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative

CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;

CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and

CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal

CC and parasitic infections.

CC Note: The sequence data for this patent did not form part of the

CC printed specification, but was obtained in electronic format directly

CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX CC

SQ Sequence 374 AA;

Query Match 100.0%; Score 1115; DB 23; Length 374;

Best Local Similarity 100.0%; Pred. No. 3.5e-105;

Matches 206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGRLLSGQPVCRGGTQPCYKVIYFHDTSRRLLNFEEAKEACRRDGGQLVSI ESEDEQKL 60

Db 22 ATGRLLSGQPVCRGGTQPCYKVIYFHDTSRRLLNFEEAKEACRRDGGQLVSI ESEDEQKL 81

QY 61 IEKFIENLLPSDGFWIGLRRREEKQSNSTACQDLYAWTDGSI SQFRNWWYVDEPSCGSEV 120

Db 82 IEKFIENLLPSDGFWIGLRRREEKQSNSTACQDLYAWTDGSI SQFRNWWYVDEPSCGSEV 141

QY 121 CVVMYHQPSAPAGIGGPPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGEETLTPV 180

Db 142 CVVMYHQPSAPAGIGGPPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGEETLTPV 201

QY 181 LPEETQEEADAKKTFKESREAAALNLAY 206

Db 202 LPEETQEEADAKKTFKESREAAALNLAY 227

RESULT 5

AAM25796

ID AAM25796 standard; Protein; 387 AA.

XX AC AAM25796;

XX 16-OCT-2001 (first entry)

XX DE Human protein sequence SEQ ID NO:1311.

XX KW Human; cancer; ulcer; HIV infection; human immunodeficiency virus;

KW antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;

KW antibacterial; endocrine; cardiant; central nervous system; virucide;

KW anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;

KW antiaggregant; haemostatic; vulnery; antiulcer; osteopathic; eczema;

KW dermatological; antiallergic; antiasthmatic; antidiabetic; cytostatic;

KW neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;

KW immunostimulant; gene therapy; antisenese therapy; vaccine; inflammation;

KW antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;

KW cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;

KW genetic disease; haematopoietic disorder; platelet disorder; asthma;

KW thrombocytopaenia; osteoporosis; severe combined immunodeficiency;

KW allergic rhinitis; diabetes; multiple sclerosis; depression;

KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;

XX KW neurological disorder.

OS Homo sapiens.

XX

PN WO200153455-A2.

XX 26-JUL-2001.

PD

XX 22-DEC-2000; 2000WO-US35017.

PF

XX 23-DEC-1999; 99US-0471275.

PR 21-JAN-2000; 2000US-0488725.

PR 25-APR-2000; 2000US-0552317.

XX (HYSE-) HYSEQ INC.

PA

XX Tang YT, Liu C, Drmanac RT;

PI WPI; 2001-457603/49.

DR N-PSDB; AAH99737.

DR

XX Isolated human polynucleotides encoding polypeptides, useful for the

PT treatment and diagnosis of e.g. cancer, ulcers and HIV infection -

PT

XX Claim 20; Page 272; 1217pp; English.

PS

XX

CC AAH99166 to AAH99904 encode the human proteins given in AAM25225 to

CC AAM25963. The proteins can have activities based on the tissues and

CC cells they are expressed in, such as: antiinflammatory; antirheumatic;

CC antiarthritic; immunosuppressive; antibacterial; endocrine; cardiant;

CC central nervous system; virucide; anti-HIV; fungicide; antimutagen;

CC cardiovascular; antianaemic; antiaggregant; haemostatic; vulnery;

CC antiulcer; osteopathic; dermatological; antiallergic; antiasthmatic;

CC antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;

CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides

CC encoding them can be used in gene therapy, antisenese therapy and vaccine

CC production. The proteins and polynucleotides are useful for screening for

CC agonists or antagonists of a protein and for the treatment and diagnosis

CC of disorders associated with the activity of a protein e.g. inflammation,

CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,

CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal

CC infections, autoimmunity, genetic diseases, haematopoietic disorders,

CC anaemia, platelet disorders, thrombocytopaenia, wounds, ulcers,

CC osteoporosis, severe combined immunodeficiency, eczema, allergic

CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,

CC Alzheimer's disease, Parkinson's disease, neurodegenerative and

CC neurological disorders.

XX

SQ Sequence 387 AA;

Query Match 100.0%; Score 1115; DB 22; Length 387;

Best Local Similarity 100.0%; Pred. No. 3.6e-105;

Matches 206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGRLLSGQPVCRGGTQPCYKVIYFHDTSRRLLNFEEAKEACRRDGGQLVSI ESEDEQKL 60

Db 35 ATGRLLSGQPVCRGGTQPCYKVIYFHDTSRRLLNFEEAKEACRRDGGQLVSI ESEDEQKL 94

QY 61 IEKFIENLLPSDGFWIGLRRREEKQSNSTACQDLYAWTDGSI SQFRNWWYVDEPSCGSEV 120

Db 95 IEKFIENLLPSDGFWIGLRRREEKQSNSTACQDLYAWTDGSI SQFRNWWYVDEPSCGSEV 154

QY 121 CVVMYHQPSAPAGIGGPPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGEETLTPV 180

Db 155 CVVMYHQPSAPAGIGGPPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGEETLTPV 214

QY 181 LPEETQEEADAKKTFKESREAAALNLAY 206

Db 215 LPEETQEEADAKKTFKESREAAALNLAY 240

RESULT 6

AA91490

ID AA91490 standard; Protein; 374 AA.

XX

AC AA91490;

XX

DT 29-JUN-2000 (first entry)
DE Human secreted protein sequence encoded by gene 40 SEQ ID NO:163.
XX
KW Human; secreted protein; diagnosis; cytostatic; immunosuppressive;
KW antiHIV; antiinflammatory; nootropic; neuroprotective; antiallergic;
KW osteopathic; antiarthritic; antibacterial; antidiabetic; antiasthma;
KW antipsoriatic; cardiant; gene therapy; cancer; neurological disorder;
KW immune disease; inflammation; blood disorder; tumour.
XX
OS Homo sapiens.
XX
XX WO200006698-A1.
PN
XX 10-FEB-2000.
PD
XX
XX 29-JUL-1999; 99WO-US17130.
XX
XX 30-JUL-1998; 98US-0094657.
PR
XX 05-AUG-1998; 98US-0095486.
PR
XX 06-AUG-1998; 98US-0095454.
PR
XX 06-AUG-1998; 98US-0095455.
PR
XX 12-AUG-1998; 98US-0096319.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX
XX Komatsoulis GA, Rosen CA, Ruben SM, Duan R, Moore PA, Shi Y;
PI Lafleur D, Wei Y, Ni J, Florence KA, Young PE, Brewer LA;
PI Soppet DR, Endress GA, Ebner R, Olsen HS, Mucenski M;
XX
DR WPI; 2000-195282/17.
DR N-PSDB; AAA26385.
XX
PT New isolated human genes and the secreted polypeptides they encode,
PT useful for diagnosis and treatment of e.g. cancers, neurological
PT disorders, immune diseases, inflammation or blood disorders
XX
PS Claim 11; Page 483-484; 634pp; English.
XX
CC The polynucleotide sequences given in AAA26346 to AAA26458 encode the
CC human secreted proteins given in AAY91451 to AAY91691. The human secreted
CC proteins can have activities based on the tissues and cells they are
CC expressed in. Examples of the activities are: cytostatic;
CC immunosuppressive; antiHIV; antiinflammatory; nootropic; neuroprotective;
CC antiallergic; osteopathic; antiarthritic; antibacterial; antidiabetic;
CC antiasthma; antipsoriatic; and cardiant. The polynucleotides and their
CC corresponding secreted proteins are useful for preventing, treating or
CC ameliorating medical conditions, e.g. by protein or gene therapy. Also
CC pathological conditions can be diagnosed by determining the amount of the
CC proteins in a sample or by determining the presence of mutations in the
CC polynucleotides. Specific uses are described for each of the
CC polynucleotides, based on which tissues they are most highly expressed
CC in, and include developing products for the diagnosis or treatment of
CC cancer, tumours, neurodegenerative disorders, developmental abnormalities
CC and foetal deficiencies, blood disorders, diseases of the immune system,
CC auto-immune diseases, hepatic and renal disease, inflammation,
CC allergies, Alzheimer's and behavioural disorders, schizophrenia,
CC osteoporosis, arthritis, infections, AIDS, spinal cord injuries,
CC transplant rejection, diabetes, asthma, sepsis, acne, psoriasis,
CC cardiovascular disorders, reproductive disorders, gastrointestinal
CC disorders, respiratory disorders and metabolic disorders. The proteins
CC or polynucleotides can also be used as food additives or preservatives.
CC The proteins are also useful for identifying their binding partners.
CC AAA26337 to AAA26345 and AAY91450 are sequences used in the
CC exemplification of the present invention.
XX
SQ Sequence 374 AA;

Query Match 99.1%; Score 1105; DB 21; Length 374;
Best Local Similarity 99.0%; Pred. No. 3.7e-104;
Matches 204; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGRLLSGQPVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI ESEDEQKL 60

Db 22 ATGRLLSGQPVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI ESEDEQKL 81
QY 61 IEKFIENLLPSDGFWIGLRRRREEKQSNSTACQDLYAWTDGSI SQFRNMYVDEPSCGSEV 120
Db 82 IEXFIENLLPSDGFWIGLRRRREEKQSNSTXCDLYAWTDGSI SQFRNMYVDEPSCGSEV 141
QY 121 CVVMYHQPSAPAGIGGYPYMFQWNDRCNMKNFICKYSDEKPAVPSREAEGEETELTPV 180
Db 142 CVVMYHQPSAPAGIGGYPYMFQWNDRCNMKNFICKYSDEKPAVPSREAEGEETELTPV 201
QY 181 LPEETQEEDAKKTFKESREAAALNLAY 206
Db 202 LPEETQEEDAKKTFKESREAAALNLAY 227

RESULT 7
AAY13367
ID AAY13367 standard; Protein; 382 AA.
XX
AC AAY13367;
XX
DT 25-JUN-1999 (first entry)
XX
DE Amino acid sequence of protein PRO2334.
XX
KW Secreted protein; transmembrane protein; human; enterocolitis;
KW Zollinger-Ellison syndrome; gastrointestinal ulceration;
KW congenital microvillus atrophy; skin disease; cell growth;
KW abnormal keratinocyte differentiation; psoriasis; epithelial cancer;
KW Parkinson's disease; Alzheimer's disease; ALS; neuropathy;
KW fibromodulin; dermal scarring; Usher Syndrome; Atrophia areata;
KW anti-thrombotic; wound healing; tissue repair.
XX
OS Homo sapiens.
XX
PN WO9914328-A2.
XX
PD 25-MAR-1999.
XX
PF 16-SEP-1998; 98WO-US19330.
XX
XX 25-NOV-1997; 97US-0066840.
PR 17-SEP-1997; 97US-0059113.
PR 17-SEP-1997; 97US-0059115.
PR 17-SEP-1997; 97US-0059117.
PR 17-SEP-1997; 97US-0059119.
PR 17-SEP-1997; 97US-0059121.
PR 17-SEP-1997; 97US-0059122.
PR 17-SEP-1997; 97US-0059184.
PR 18-SEP-1997; 97US-0059263.
PR 18-SEP-1997; 97US-0059266.
PR 15-OCT-1997; 97US-0062125.
PR 17-OCT-1997; 97US-0062285.
PR 17-OCT-1997; 97US-0062287.
PR 21-OCT-1997; 97US-0063486.
PR 24-OCT-1997; 97US-0062814.
PR 24-OCT-1997; 97US-0062816.
PR 24-OCT-1997; 97US-0063045.
PR 24-OCT-1997; 97US-0063120.
PR 24-OCT-1997; 97US-0063121.
PR 24-OCT-1997; 97US-0063127.
PR 24-OCT-1997; 97US-0063128.
PR 27-OCT-1997; 97US-0063329.
PR 27-OCT-1997; 97US-0063327.
PR 28-OCT-1997; 97US-0063541.
PR 28-OCT-1997; 97US-0063542.
PR 28-OCT-1997; 97US-0063544.
PR 28-OCT-1997; 97US-0063549.
PR 28-OCT-1997; 97US-0063550.
PR 28-OCT-1997; 97US-0063564.
PR 29-OCT-1997; 97US-0063435.
PR 29-OCT-1997; 97US-0063704.

PR 29-OCT-1997; 97US-0063732.
PR 29-OCT-1997; 97US-0063738.
PR 29-OCT-1997; 97US-0063734.
PR 29-OCT-1997; 97US-0064215.
PR 29-OCT-1997; 97US-0063735.
PR 31-OCT-1997; 97US-0063870.
PR 31-OCT-1997; 97US-0064103.
PR 03-NOV-1997; 97US-0064248.
PR 07-NOV-1997; 97US-0064809.
PR 12-NOV-1997; 97US-0065186.
PR 17-NOV-1997; 97US-0065846.
PR 18-NOV-1997; 97US-0065693.
PR 21-NOV-1997; 97US-0066120.
PR 21-NOV-1997; 97US-0066364.
PR 24-NOV-1997; 97US-0066772.
PR 24-NOV-1997; 97US-0066466.
PR 24-NOV-1997; 97US-0066770.
PR 24-NOV-1997; 97US-0066511.
PR 24-NOV-1997; 97US-0066453.
XX
PA (GETH) GENENTECH INC.
XX
PI Chen J, Goddard A, Gurney AL, Pennica D, Wood WI, Yuan J;
XX
DR WPI; 1999-229533/19.
DR N-PSDB; AAX52238.
XX
PT New isolated human genes and polypeptides used in, e.g. treatment of
PT gastrointestinal ulceration
XX
PS Claim 12; Fig 50; 320pp; English.
XX
CC AAY13344-403 represent secreted and transmembrane human proteins.
CC The cDNA sequences are obtained from cDNA libraries, prepared from
CC fetal lung, fetal kidney, fetal brain, fetal liver and fetal retina.
CC The encoded polypeptides have specific uses based on their homology to
CC known polypeptides, e.g. PRO211 and PRO217 can be used for disorders
CC associated with the preservation and maintenance of gastrointestinal
CC mucosa and the repair of acute and chronic mucosal lesions
CC (e.g. enterocolitis, Zollinger-Ellison syndrome, gastrointestinal
CC ulceration and congenital microvillus atrophy), skin diseases associated
CC with abnormal keratinocyte differentiation (e.g. psoriasis, epithelial
CC cancers such as lung squamous cell carcinoma of the vulva and gliomas),
CC potent effects on cell growth and development, diseases related to
CC growth or survival of nerve cells including Parkinson's disease,
CC Alzheimer's disease, ALS, neuropathies or cancer. PRO265 can be used as
CC for fibromodulin, e.g. for reducing dermal scarring. PRO264 can be used
CC as a target for anti-tumor drugs. PRO533 may be used in the treatment
CC of Usher Syndrome or Atrophia areata; PRO269 can be used as an
CC anti-thrombotic agent; PRO287 polypeptides and portions may have
CC therapeutic applications in wound healing and tissue repair; PRO317 can
CC be used for treating problems of the kidney, uterus, endometrium, blood
CC vessels, or related tissue, e.g. in the heart of genital tract.
XX
SQ Sequence 382 AA;

Query Match 98.7%; Score 1101; DB 20; Length 382;
Best Local Similarity 96.3%; Pred. No. 9.6e-104;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 ATGRLLS-----GQPVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
| | | | |
DB 22 ATGRLLSASDLRLGGQPVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81
| | | | |
QY 53 ESEDEQKLIETIENLLPSDGFWIGLRRREKQSNSTACQDLYAWNTDGSISQFRNYYVD 112
| | | | |
DB 82 ESEDEQKLIETIENLLPSDGFWIGLRRREKQSNSTACQDLYAWNTDGSISQFRNYYVD 141
| | | | |
QY 113 EPSCGSEVCVMYHQPSAPAGIGGYPYMFQWDDRCNMKNFICKYSDEKPAVPSRAEAGE 172
| | | | |
DB 142 EPSCGSEVCVMYHQPSAPAGIGGYPYMFQWDDRCNMKNFICKYSDEKPAVPSRAEAGE 201
| | | | |
QY 173 ETELTPVLPETQEDAKKTFKESREAAALNLAY 206

Db 202 ETELTPVLPETQEDAKKTFKESREAAALNLAY 235
| | | | |
RESULT 8
AAU29033
ID AAU29033 standard; Protein; 382 AA.
XX
AC AAU29033;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human PRO polypeptide sequence #10.
XX
KW PRO polypeptide; mammal; tumour; cancer; human; cattle; horse; sheep;
KW dog; cat; pig; goat; rabbit; tumour necrosis factor alpha; TNF-alpha;
KW blood; chondrocyte cell; cell proliferation; cell differentiation; colon;
KW adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200168848-A2.
XX
PD 20-SEP-2001.
XX
PF 28-FEB-2001; 2001WO-US06520.
XX
PR 01-MAR-2000; 2000WO-US05601.
PR 02-MAR-2000; 2000WO-US05841.
PR 03-MAR-2000; 2000US-187202P.
PR 06-MAR-2000; 2000US-186968P.
PR 14-MAR-2000; 2000US-189320P.
PR 14-MAR-2000; 2000US-189328P.
PR 15-MAR-2000; 2000WO-US06884.
PR 21-MAR-2000; 2000US-190828P.
PR 21-MAR-2000; 2000US-191007P.
PR 21-MAR-2000; 2000US-191048P.
PR 21-MAR-2000; 2000US-191314P.
PR 28-MAR-2000; 2000US-192655P.
PR 29-MAR-2000; 2000US-193032P.
PR 29-MAR-2000; 2000US-193053P.
PR 30-MAR-2000; 2000WO-US08439.
PR 04-APR-2000; 2000US-194449P.
PR 04-APR-2000; 2000US-194647P.
PR 11-APR-2000; 2000US-195975P.
PR 11-APR-2000; 2000US-196000P.
PR 11-APR-2000; 2000US-196187P.
PR 11-APR-2000; 2000US-196690P.
PR 11-APR-2000; 2000US-196820P.
PR 18-APR-2000; 2000US-198121P.
PR 18-APR-2000; 2000US-198585P.
PR 25-APR-2000; 2000US-199397P.
PR 25-APR-2000; 2000US-199550P.
PR 25-APR-2000; 2000US-199654P.
PR 03-MAY-2000; 2000US-201516P.
PR 17-MAY-2000; 2000WO-US13705.
PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.
PR 02-JUN-2000; 2000WO-US15264.
PR 05-JUN-2000; 2000US-209832P.
PR 28-JUL-2000; 2000WO-US20710.
PR 22-AUG-2000; 2000US-0644848.
PR 24-AUG-2000; 2000WO-US23328.
PR 08-NOV-2000; 2000WO-US30952.
PR 01-DEC-2000; 2000WO-US32678.
PR 20-DEC-2000; 2000WO-US34956.
XX
PA (GETH) GENENTECH INC.
XX
PI Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;
PI Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;
XX
DR WPI; 2001-602746/68.

DR N-PSDB; AAS45934.
XX Novel nucleic acids encoding PRO polypeptides, used to diagnose the
PT presence of tumours, such as prostate and breast tumours, in mammals and
PT to screen for modulators of the compounds -
XX Claim 11; Fig 20; 774pp; English.
XX Sequences AAU29024-AAU29328 represent PRO polypeptides of the invention.
CC The PRO polypeptides and their associated nucleic acids can be used to
CC detect the presence of a tumour in a mammal by comparing the level of
CC expression of a PRO polypeptide in a test sample of cells from the animal
CC and a control sample of normal cells, whereby a higher level of
CC expression in the test sample indicates the presence of a tumour in the
CC mammal. Mammals include dogs, cats, cattle, horses, sheep, pigs, goats
CC and rabbits but are preferably human. The polypeptides can be used to
CC stimulate tumour necrosis factor (TNF) alpha release from human blood,
CC when contacted with it. A specific polypeptide can be used to stimulate
CC the proliferation or differentiation of chondrocyte cells. The PRO
CC proteins can be used to determine the presence of tumours and also
CC susceptibility to tumour development, particularly adrenal, lung, colon,
CC breast, prostate, rectal, cervical, or liver tumours, in mammalian
CC subjects. The oligonucleotide probes specific for the PRO nucleic acids
CC can be used for genetic analysis of individuals with genetic disorders.
XX Sequence 382 AA;

Query Match 98.7%; Score 1101; DB 22; Length 382;
Best Local Similarity 96.3%; Pred. No. 9.6e-104;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
QY 1 ATGRLLS-----GQVCRGGTQPCYKVIYFHDTSRRLNFEAEKACRRDGGQLVSI 52
Db 22 ATGRLLSASDLDRGGQPCVCRGGTQPCYKVIYFHDTSRRLNFEAEKACRRDGGQLVSI 81
QY 53 ESEDEQKLIKFIENLLPSDGFWIGLRRRREKQSNSTACQDLYAWTDGSIQFRNWWYVD 112
Db 82 ESEDEQKLIKFIENLLPSDGFWIGLRRRREKQSNSTACQDLYAWTDGSIQFRNWWYVD 141
QY 113 EPSCGSEVCVMYHQPAPAGIGGYPYFQWDDRCNMKNFICKYSDEKPAVPSREAEGE 172
Db 142 EPSCGSEVCVMYHQPAPAGIGGYPYFQWDDRCNMKNFICKYSDEKPAVPSREAEGE 201
QY 173 ETELTPVLPEETQEEDAKKTFKESREAAALNLAY 206
Db 202 ETELTPVLPEETQEEDAKKTFKESREAAALNLAY 235

RESULT 9
AAB80235
ID AAB80235 standard; Protein; 382 AA.
XX AAB80235;
AC AAB80235;
XX 24-APR-2001 (first entry)
DT Human PRO234 protein.
DE Human; PRO; dermatological; antipsoriatic; cytostatic; antiinflammatory;
KW antiparkinsonian nootropic; neuroprotective; vulnery; cardiant;
KW antiangiogenic; vasotropic; antiasthmatic; antirheumatic; cancer;
KW antiarthritic; antinfertility; antidiabetic; antiviral; diabetes;
KW ophthalmological; gene therapy; skin disease; gastrointestinal disorder;
KW ischaemia; inflammation.

XX Homo sapiens.
OS WO200104311-A1.
XX 18-JAN-2001.
PD 22-FEB-2000; 2000WO-US04414.
PF 22-FEB-2000; 2000WO-US04414.
XX

PR 07-JUL-1999; 99US-0143048.
PR 26-JUL-1999; 99US-0145698.
PR 28-JUL-1999; 99US-0146222.
PR 08-SEP-1999; 99WO-US20594.
PR 13-SEP-1999; 99WO-US20944.
PR 15-SEP-1999; 99WO-US21090.
PR 15-SEP-1999; 99WO-US21547.
PR 05-OCT-1999; 99WO-US23089.
PR 29-NOV-1999; 99WO-US28214.
PR 30-NOV-1999; 99WO-US28313.
PR 16-DEC-1999; 99WO-US30095.
PR 20-DEC-1999; 99WO-US30911.
PR 20-DEC-1999; 99WO-US30999.
PR 05-JAN-2000; 99WO-US00219.
XX (GETH) GENENTECH INC.
PA Ashkenazi AJ, Botstein D, Desnoyers L, Eaton DL, Ferrara N;
XX Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME, Goddard A;
PI Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ, Kljavin IJ;
PI Mather JP, Pan J, Paoni NF, Roy MA, Stewart TA, Tumas D;
PI Williams PM, Wood WI;
XX WPI; 2001-081051/09.
DR N-PSDB; AAF72396.
XX

Sixty one nucleic acids encoding PRO polypeptides which are useful in
PT the treatment of skin diseases (e.g. psoriasis), cancers (e.g. lung
PT squamous cell carcinoma) and neurodegenerative diseases (e.g.
PT Alzheimer's disease) -

Claim 1; Fig 50; 393pp; English.

The present sequence is one of sixty one novel secreted and
CC transmembrane PRO polypeptides. The PRO polypeptides are
CC useful for treating skin diseases (e.g. psoriasis), cancers (e.g. lung
CC squamous cell carcinoma), gastrointestinal disorders (e.g.
CC enterocolitis), neurodegenerative diseases (e.g. Alzheimer's disease,
CC Parkinson's disease), wound repair, cardiovascular disorders (e.g.
CC endometrial bleeding angiogenesis, ischaemias such as coronary
CC ischaemia, atherosclerosis), inflammatory disorders (e.g. asthma,
CC rheumatoid arthritis, multiple sclerosis), infertility, AIDS and
CC diabetes and retinal disorders such as retinitis pigmentosum.
CC The PRO nucleic acids have applications in molecular biology, including
CC use as hybridization probes, and in chromosome and gene mapping.

SQ Sequence 382 AA;

Query Match 98.7%; Score 1101; DB 22; Length 382;
Best Local Similarity 96.3%; Pred. No. 9.6e-104;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 ATGRLLS-----GQVCRGGTQPCYKVIYFHDTSRRLNFEAEKACRRDGGQLVSI 52
Db 22 ATGRLLSASDLDRGGQPCVCRGGTQPCYKVIYFHDTSRRLNFEAEKACRRDGGQLVSI 81
QY 53 ESEDEQKLIKFIENLLPSDGFWIGLRRRREKQSNSTACQDLYAWTDGSIQFRNWWYVD 112
Db 82 ESEDEQKLIKFIENLLPSDGFWIGLRRRREKQSNSTACQDLYAWTDGSIQFRNWWYVD 141
QY 113 EPSCGSEVCVMYHQPAPAGIGGYPYFQWDDRCNMKNFICKYSDEKPAVPSREAEGE 172
Db 142 EPSCGSEVCVMYHQPAPAGIGGYPYFQWDDRCNMKNFICKYSDEKPAVPSREAEGE 201
QY 173 ETELTPVLPEETQEEDAKKTFKESREAAALNLAY 206
Db 202 ETELTPVLPEETQEEDAKKTFKESREAAALNLAY 235

RESULT 10
ABU69645
ID ABU69645 standard; Protein; 382 AA.
XX

AC ABU69645;
XX
DT 05-JUN-2003 (first entry)
XX
DE Novel human secreted and transmembrane protein PRO234.
XX
KW Human; secreted and transmembrane protein; gene therapy; psoriasis;
KW enterocolitis; gastrointestinal ulceration; skin disease;
KW keratinocyte differentiation; epithelial cancer; Alzheimer's disease;
KW squamous cell carcinoma; Parkinson's disease; inflammatory disease;
KW amyotrophic lateral sclerosis; rheumatoid arthritis; asthma;
KW multiple sclerosis; organ failure; atherosclerosis; cardiac injury;
KW infertility; birth defect; premature aging; AIDS; cancer;
KW diabetic complication; wound repair; tissue re-growth.
XX
OS Homo sapiens.
XX
PN US2003017463-A1.
XX
PD 23-JAN-2003.
XX
PF 11-JUL-2001; 2001US-0903640.
XX
PR 10-SEP-1998; 98WO-US18824.
PR 14-SEP-1998; 98WO-US19177.
PR 16-SEP-1998; 98WO-US19330.
PR 17-SEP-1998; 98WO-US19437.
PR 01-DEC-1998; 98WO-US25108.
PR 08-SEP-1999; 99WO-US20594.
PR 13-SEP-1999; 99WO-US20944.
PR 15-SEP-1999; 99WO-US21547.
PR 05-OCT-1999; 99WO-US23089.
PR 29-NOV-1999; 99WO-US28214.
PR 30-NOV-1999; 99WO-US28313.
PR 01-DEC-1999; 99WO-US28301.
PR 02-DEC-1999; 99WO-US28564.
PR 02-DEC-1999; 99WO-US28565.
PR 16-DEC-1999; 99WO-US30095.
PR 20-DEC-1999; 99WO-US30911.
PR 20-DEC-1999; 99WO-US30999.
PR 05-JAN-2000; 2000WO-US00219.
PR 11-FEB-2000; 2000WO-US03565.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US05004.
PR 02-MAR-2000; 2000WO-US05841.
PR 20-MAR-2000; 2000WO-US07377.
PR 30-MAR-2000; 2000WO-US08439.
PR 22-MAY-2000; 2000WO-US14042.
PR 02-JUN-2000; 2000WO-US15264.
PR 28-JUL-2000; 2000WO-US20710.
PR 24-AUG-2000; 2000WO-US23328.
PR 17-SEP-1997; 97US-059113P.
PR 17-SEP-1997; 97US-059115P.
PR 17-SEP-1997; 97US-059117P.
PR 17-SEP-1997; 97US-059119P.
PR 17-SEP-1997; 97US-059121P.
PR 17-SEP-1997; 97US-059122P.
PR 17-SEP-1997; 97US-059184P.
PR 18-SEP-1997; 97US-059263P.
PR 18-SEP-1997; 97US-059266P.
PR 15-OCT-1997; 97US-062125P.
PR 17-OCT-1997; 97US-062285P.
PR 17-OCT-1997; 97US-062287P.
PR 21-OCT-1997; 97US-063486P.
PR 24-OCT-1997; 97US-062814P.
PR 24-OCT-1997; 97US-062816P.
PR 24-OCT-1997; 97US-063045P.
PR 24-OCT-1997; 97US-063120P.
PR 24-OCT-1997; 97US-063121P.
PR 24-OCT-1997; 97US-063127P.
PR 24-OCT-1997; 97US-063128P.
PR 27-OCT-1997; 97US-063327P.

PR 27-OCT-1997; 97US-063329P.
PR 28-OCT-1997; 97US-063541P.
PR 28-OCT-1997; 97US-063542P.
PR 28-OCT-1997; 97US-063544P.
PR 28-OCT-1997; 97US-063549P.
PR 28-OCT-1997; 97US-063550P.
PR 28-OCT-1997; 97US-063564P.
PR 29-OCT-1997; 97US-063435P.
PR 29-OCT-1997; 97US-063704P.
PR 29-OCT-1997; 97US-063732P.
PR 29-OCT-1997; 97US-063734P.
PR 29-OCT-1997; 97US-063735P.
PR 29-OCT-1997; 97US-063738P.
PR 29-OCT-1997; 97US-064215P.
PR 31-OCT-1997; 97US-063870P.
PR 31-OCT-1997; 97US-064103P.
PR 03-NOV-1997; 97US-064248P.
PR 07-NOV-1997; 97US-064809P.
PR 12-NOV-1997; 97US-065186P.
PR 17-NOV-1997; 97US-065846P.
PR 18-NOV-1997; 97US-065693P.
PR 21-NOV-1997; 97US-066120P.
PR 21-NOV-1997; 97US-066364P.
PR 24-NOV-1997; 97US-066453P.
PR 24-NOV-1997; 97US-066466P.
PR 24-NOV-1997; 97US-066511P.
PR 24-NOV-1997; 97US-066770P.
PR 24-NOV-1997; 97US-066772P.
PR 25-NOV-1997; 97US-066840P.
PR 12-DEC-1997; 97US-069425P.
PR 04-JUN-1998; 98US-088026P.
PR 10-SEP-1998; 98US-099803P.
PR 14-SEP-1998; 98US-100262P.
PR 17-SEP-1998; 98US-100858P.
PR 13-OCT-1998; 98US-104080P.
PR 20-NOV-1998; 98US-109304P.
PR 22-DEC-1998; 98US-113296P.
PR 07-JUL-1999; 99US-143048P.
PR 26-JUL-1999; 99US-145698P.
PR 28-JUL-1999; 99US-146222P.
PR 18-SEP-2000; 2000US-0665350.
XX
PA (GETH) GENENTECH INC.

Ashkenazi A, Botstein D, Desnoyers L, Eaton DL, Ferrara N;
Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME, Goddard A;
Godowski PJ, Grimaldi JC, Gurney AL, Hillan KJ, Kljavin IJ;
Mather JP, Pan J, Paoni NF, Roy MA, Stewart TA, Tumas D;
Williams PM, Wood WI;

WPI; 2003-341586/32.
N-PSDB; ACA54920.

New PRO polypeptides and nucleic acid molecules, useful in diagnosing
or treating inflammatory diseases, organ failure, atherosclerosis,
cardiac injury, infertility, cancer, AIDS, Alzheimer's disease or
Parkinson's disease -

Claim 12; Fig 50; 473pp; English.

The invention describes sixty one nucleic acids encoding PRO polypeptides
(secreted and transmembrane). The PRO polypeptides and nucleic acids are
useful in diagnosing or treating enterocolitis, gastrointestinal
ulceration, skin diseases associated with abnormal keratinocyte
differentiation, e.g. psoriasis or epithelial cancers such as squamous
cell carcinoma, Alzheimer's disease, Parkinson's disease, amyotrophic
lateral sclerosis, inflammatory diseases, e.g. rheumatoid arthritis,
asthma or multiple sclerosis, organ failure, atherosclerosis, cardiac
injury, infertility, birth defects, premature aging, AIDS, cancer,
diabetic complications, or mutations in general. The polypeptides are
also useful for wound repair and associated therapies concerned with
re-growth of tissue. The PRO polypeptides and nucleic acid molecules
are also useful in gene therapy, and as molecular weight markers for

CC protein electrophoresis purposes. The anti-PRO antibodies may be used
CC in diagnostic assays for PRO, or for the affinity purification of PRO
CC from recombinant cell culture or natural sources. This is the amino
CC acid sequence of a novel human PRO polypeptide.
XX
SQ Sequence 382 AA;

Query Match 98.7%; Score 1101; DB 24; Length 382;
Best local Similarity 96.3%; Pred. No. 9.6e-104;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps .1;

QY 1 ATGRLLS-----GQVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
| | | | |
Db 22 ATGRLLSASDLRGGQPVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

QY 53 ESEDEQKLIKFIENLLPSDGFWIGLRRREKQSNSTACQDLYAWTDGSIQFRNWWYVD 112
| | | | |
Db 82 ESEDEQKLIKFIENLLPSDGFWIGLRRREKQSNSTACQDLYAWTDGSIQFRNWWYVD 141

QY 113 EPSCGSEVCVMYHQPSAPAGIGGYPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGE 172
| | | | |
Db 142 EPSCGSEVCVMYHQPSAPAGIGGYPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGE 201

QY 173 ETELTTPVLPEETQEDAKKTFKESREAAINLAY 206
| | | | |
Db 202 ETELTTPVLPEETQEDAKKTFKESREAAINLAY 235

RESULT 11
ABU71121
ID ABU71121 standard; Protein; 382 AA.
XX
AC ABU71121;
XX
DT 10-JUN-2003 (first entry)
XX
DE Human PRO234 protein.
XX
KW Human; PRO; secreted; transmembrane; cytostatic; TNF-alpha; blood;
KW tumour necrosis factor alpha release; chondrocyte cell; proliferation;
KW differentiation; tumour; gene therapy.
XX
OS Homo sapiens.
XX
PN US2003036143-A1.
XX
PD 20-FEB-2003.
XX
PF 02-JUL-2002; 2002US-0187600.
XX
PR 16-SEP-1998; 98WO-US19330.
PR 07-OCT-1998; 98WO-US21141.
PR 01-DEC-1998; 98WO-US25108.
PR 08-MAR-1999; 99WO-US05028.
PR 14-MAY-1999; 99WO-US10733.
PR 02-JUN-1999; 99WO-US12252.
PR 01-SEP-1999; 99WO-US20111.
PR 15-SEP-1999; 99WO-US21090.
PR 01-DEC-1999; 99WO-US28301.
PR 02-DEC-1999; 99WO-US28551.
PR 30-DEC-1999; 99WO-US31274.
PR 05-JAN-2000; 2000WO-US00219.
PR 18-FEB-2000; 2000WO-US04341.
PR 18-FEB-2000; 2000WO-US04342.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US05004.
PR 01-MAR-2000; 2000WO-US05601.
PR 02-MAR-2000; 2000WO-US05841.
PR 15-MAR-2000; 2000WO-US06884.
PR 30-MAR-2000; 2000WO-US08439.
PR 17-MAY-2000; 2000WO-US13705.
PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.

PR 02-JUN-2000; 2000WO-US15264.
PR 28-JUL-2000; 2000WO-US20710.
PR 24-AUG-2000; 2000WO-US23328.
PR 08-NOV-2000; 2000WO-US30952.
PR 01-DEC-2000; 2000WO-US32678.
PR 20-DEC-2000; 2000WO-US34956.
PR 28-FEB-2001; 2001WO-US06520.
PR 01-JUN-2001; 2001WO-US17800.
PR 20-JUN-2001; 2001WO-US19692.
PR 29-JUN-2001; 2001WO-US21066.
PR 09-JUL-2001; 2001WO-US21735.
PR 29-AUG-2001; 2001WO-US27099.
PR 18-SEP-1997; 97US-059263P.
PR 18-SEP-1997; 97US-059266P.
PR 17-OCT-1997; 97US-062250P.
PR 21-OCT-1997; 97US-063486P.
PR 24-OCT-1997; 97US-063120P.
PR 24-OCT-1997; 97US-063121P.
PR 28-OCT-1997; 97US-063540P.
PR 28-OCT-1997; 97US-063541P.
PR 28-OCT-1997; 97US-063544P.
PR 28-OCT-1997; 97US-063564P.
PR 29-OCT-1997; 97US-063734P.
PR 31-OCT-1997; 97US-063870P.
PR 31-OCT-1997; 97US-064103P.
PR 13-NOV-1997; 97US-065311P.
PR 21-NOV-1997; 97US-066120P.
PR 24-NOV-1997; 97US-066466P.
PR 24-NOV-1997; 97US-066772P.
PR 11-DEC-1997; 97US-069335P.
PR 12-DEC-1997; 97US-069425P.
PR 17-DEC-1997; 97US-069870P.
PR 18-DEC-1997; 97US-068017P.
PR 10-MAR-1998; 98US-077450P.
PR 11-MAR-1998; 98US-077632P.
PR 11-MAR-1998; 98US-077649P.
PR 20-MAR-1998; 98US-078886P.
PR 20-MAR-1998; 98US-078939P.
PR 27-MAR-1998; 98US-079664P.
PR 27-MAR-1998; 98US-079786P.
PR 31-MAR-1998; 98US-080107P.
PR 31-MAR-1998; 98US-080194P.
PR 01-APR-1998; 98US-080327P.
PR 01-APR-1998; 98US-080333P.
PR 08-APR-1998; 98US-081049P.
PR 08-APR-1998; 98US-081070P.
PR 09-APR-1998; 98US-081195P.
PR 15-APR-1998; 98US-081838P.
PR 21-APR-1998; 98US-082568P.
PR 21-APR-1998; 98US-082569P.
PR 22-APR-1998; 98US-082704P.
PR 22-APR-1998; 98US-082797P.
PR 28-APR-1998; 98US-083322P.
PR 29-APR-1998; 98US-083495P.
PR 29-APR-1998; 98US-083496P.
PR 29-APR-1998; 98US-083499P.
PR 29-APR-1998; 98US-083559P.
PR 05-MAY-1998; 98US-084366P.
PR 06-MAY-1998; 98US-084414P.
PR 07-MAY-1998; 98US-084639P.
PR 07-MAY-1998; 98US-084640P.
PR 07-MAY-1998; 98US-084643P.
PR 15-MAY-1998; 98US-085579P.
PR 15-MAY-1998; 98US-085580P.
PR 15-MAY-1998; 98US-085582P.
PR 15-MAY-1998; 98US-085700P.
PR 18-MAY-1998; 98US-086023P.
PR 22-MAY-1998; 98US-086392P.
PR 22-MAY-1998; 98US-086486P.
PR 28-MAY-1998; 98US-087098P.
PR 28-MAY-1998; 98US-087208P.
PR 02-JUN-1998; 98US-087609P.
PR 02-JUN-1998; 98US-087759P.

PR	03-JUN-1998;	98US-087827P;
PR	04-JUN-1998;	98US-088025P;
PR	04-JUN-1998;	98US-088028P;
PR	04-JUN-1998;	98US-088029P;
PR	04-JUN-1998;	98US-088033P;
PR	04-JUN-1998;	98US-088326P;
PR	05-JUN-1998;	98US-088167P;
PR	05-JUN-1998;	98US-088202P;
PR	05-JUN-1998;	98US-088212P;
PR	05-JUN-1998;	98US-088217P;
PR	09-JUN-1998;	98US-088655P;
PR	10-JUN-1998;	98US-088722P;
PR	10-JUN-1998;	98US-088738P;
PR	10-JUN-1998;	98US-088740P;
PR	10-JUN-1998;	98US-088811P;
PR	10-JUN-1998;	98US-088824P;
PR	10-JUN-1998;	98US-088825P;
PR	10-JUN-1998;	98US-088826P;
PR	11-JUN-1998;	98US-088861P;
PR	11-JUN-1998;	98US-088863P;
PR	11-JUN-1998;	98US-088876P;
PR	12-JUN-1998;	98US-089090P;
PR	12-JUN-1998;	98US-089105P;
PR	16-JUN-1998;	98US-089512P;
PR	16-JUN-1998;	98US-089514P;
PR	17-JUN-1998;	98US-089538P;
PR	17-JUN-1998;	98US-089598P;
PR	17-JUN-1998;	98US-089653P;
PR	18-JUN-1998;	98US-089908P;
PR	19-JUN-1998;	98US-089952P;
PR	22-JUN-1998;	98US-090246P;
PR	22-JUN-1998;	98US-090252P;
PR	22-JUN-1998;	98US-090254P;
PR	24-JUN-1998;	98US-090429P;
PR	24-JUN-1998;	98US-090435P;
PR	24-JUN-1998;	98US-090444P;
PR	24-JUN-1998;	98US-090461P;
PR	24-JUN-1998;	98US-090535P;
PR	24-JUN-1998;	98US-090540P;
PR	25-JUN-1998;	98US-090676P;
PR	25-JUN-1998;	98US-090678P;
PR	25-JUN-1998;	98US-090688P;
PR	25-JUN-1998;	98US-090690P;
PR	25-JUN-1998;	98US-090694P;
PR	25-JUN-1998;	98US-090695P;
PR	25-JUN-1998;	98US-090696P;
PR	26-JUN-1998;	98US-090862P;
PR	26-JUN-1998;	98US-090863P;
PR	26-JUN-1998;	98US-091010P;
PR	01-JUL-1998;	98US-091359P;
PR	01-JUL-1998;	98US-091544P;
PR	02-JUL-1998;	98US-091478P;
PR	02-JUL-1998;	98US-091486P;
PR	02-JUL-1998;	98US-091626P;
PR	02-JUL-1998;	98US-091628P;
PR	02-JUL-1998;	98US-091632P;
PR	24-JUL-1998;	98US-094006P;
PR	04-AUG-1998;	98US-095282P;
PR	10-AUG-1998;	98US-095998P;
PR	10-AUG-1998;	98US-096012P;
PR	17-AUG-1998;	98US-096757P;
PR	17-AUG-1998;	98US-096766P;
PR	17-AUG-1998;	98US-096867P;
PR	17-AUG-1998;	98US-096891P;
PR	17-AUG-1998;	98US-096897P;
PR	18-AUG-1998;	98US-096949P;
PR	18-AUG-1998;	98US-096959P;
PR	26-AUG-1998;	98US-097022P;
PR	26-AUG-1998;	98US-097952P;
PR	26-AUG-1998;	98US-097954P;
PR	26-AUG-1998;	98US-097955P;
PR	26-AUG-1998;	98US-097971P;
PR	26-AUG-1998;	98US-097974P;

PR	26-AUG-1998;	98US-098014P.
PR	01-SEP-1998;	98US-098716P.
PR	01-SEP-1998;	98US-098723P.
PR	02-SEP-1998;	98US-098803P.
PR	02-SEP-1998;	98US-098821P.
PR	02-SEP-1998;	98US-098843P.
PR	09-SEP-1998;	98US-099602P.
PR	10-SEP-1998;	98US-099741P.
PR	10-SEP-1998;	98US-099754P.
PR	10-SEP-1998;	98US-099763P.
PR	10-SEP-1998;	98US-099812P.

Query Match			98.7%;	Score 1101;	DB 24;	Length 382;
Best Local Similarity			96.3%;	Pred. No. 9.6e-104;		
Matches 206; Conservative			0;	Mismatches 0;	Indels 8;	Gaps 1;

Qy	1	ATGRLLS-----GQPVCRRGGTORPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI	52
Dd	22	ATGRLLSASDLDRGGQPVCRGGTORPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI	81
Qy	53	ESEDEOKLIEKFIENTLLPSDGDFWIGLRREEKQSNSTACQDLYAWTDGSISQFRNWWVD	112
Dd	82	ESEDEOKLIEKFIENTLLPSDGDFWIGLRREEKQSNSTACQDLYAWTDGSISQFRNWWVD	141
Qy	113	EPSCGSEVCVMYHQSPAPAGIGGPYMFQWDDRCNMKNFNICKYSDEKPVPVSREAEGE	172
-			
Dd	142	EPSCGSEVCVMYHQSPAPAGIGGPYMFQWDDRCNMKNFNICKYSDEKPVPVSREAEGE	201
Qy	173	ETELTTPVLPEETQEEDAKTKFKESREALNLAY	206
Dd	202	ETELTTPVLPEETQEEDAKTKFKESREALNLAY	235

PR 22-FEB-2000; 2000WO-US04414.
PR 28-JUL-2000; 2000WO-US20710.
PR 24-AUG-2000; 2000WO-US23328.
PR 17-SEP-1997; 97US-059113P.
PR 17-SEP-1997; 97US-059115P.
PR 17-SEP-1997; 97US-059117P.
PR 18-SEP-1997; 97US-059266P.
PR 15-OCT-1997; 97US-062125P.
PR 17-OCT-1997; 97US-062285P.
PR 17-OCT-1997; 97US-062287P.
PR 21-OCT-1997; 97US-063486P.
PR 24-OCT-1997; 97US-062814P.
PR 24-OCT-1997; 97US-062816P.
XX
PA (GETH) GENENTECH INC.
XX
PI Ashkenazi A, Botstein D, Desnoyers L, Eaton DL, Ferrara N;
PI Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME, Goddard A;
PI Godowski PJ, Grimaldi JC, Gurney AL, Hillan KJ, Kljavin IJ;
PI Mather JP, Pan J, Paoni NF, Roy MA, Stewart TA, Tumas D;
PI Williams PM, Wood WI;
XX
DR WPI; 2003-361832/34.
DR N-PSDB; ACA58405.
XX
PT New isolated nucleic acid encoding a PRO polypeptide, e.g. PRO245 or
PT PRO1868, useful in molecular biology, chromosome and gene mapping, in
PT generating antisense RNA and DNA, and in gene therapy -
XX
PS Claim 12; Fig 50; 474pp; English.
XX
CC The present invention relates to the isolation of novel human secreted
CC and transmembrane proteins (PRO polypeptides), and the polynucleotide
CC sequences encoding them. The polynucleotide sequences are useful in
CC molecular biology, as hybridisation probes, in chromosome and gene
CC mapping, in generating antisense RNA and DNA, and in gene therapy. The
CC polynucleotide sequences may also be used in preparing PRO polypeptides
CC by recombinant techniques, and in generating either transgenic animals
CC or knock-out animals which, in turn, are useful in the development and
CC screening of therapeutically useful reagents. The PRO polypeptides or
CC their antibodies are useful in preparing a medicament for treating a
CC condition responsive to the polypeptide or antibody, such as cancer,
CC Alzheimer's disease or ischaemia, and in various diagnostic assays.
CC ABU71445-ABU71505 represent human PRO polypeptides of the invention.
XX
SQ Sequence 382 AA;

Query Match 98.7%; Score 1101; DB 24; Length 382;
Best Local Similarity 96.3%; Pred. No. 9.6e-104;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 ATGRLLS-----GQVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
| | | | |
Db 22 ATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81
| | | | |
QY 53 ESEDEQKLI EKFIENLLPSDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNWWYVD 112
| | | | |
Db 82 ESEDEQKLI EKFIENLLPSDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNWWYVD 141
| | | | |
QY 113 EPSCGSEVCVMYHQPSAPAGIGPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGE 172
| | | | |
Db 142 EPSCGSEVCVMYHQPSAPAGIGPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGE 201
| | | | |
QY 173 ETELTTPVLPEETQEEDAKTKFKESREAAALNLAY 206
| | | | |
Db 202 ETELTTPVLPEETQEEDAKTKFKESREAAALNLAY 235
| | | | |

RESULT 13
ABU71914
ID ABU71914 standard; Protein; 382 AA.
XX
AC ABU71914;

XX 12-JUN-2003 (first entry)
DT Human secreted/transmembrane protein PRO234.
XX
DE
XX
KW Human; secreted protein; transmembrane protein; PRO;
KW gene therapy; chromosome identification; chromosome marker.
XX
OS Homo sapiens.
XX
PN US2003003530-A1.
XX
PD 02-JAN-2003.
XX
PF 11-JUL-2001; 2001US-0904011.
XX
PR 10-SEP-1998; 98WO-US18824.
PR 14-SEP-1998; 98WO-US19177.
PR 16-SEP-1998; 98WO-US19330.
PR 17-SEP-1998; 98WO-US19437.
PR 01-DEC-1998; 98WO-US25108.
PR 08-SEP-1999; 99WO-US20594.
PR 13-SEP-1999; 99WO-US20944.
PR 15-SEP-1999; 99WO-US21090.
PR 15-SEP-1999; 99WO-US21547.
PR 05-OCT-1999; 99WO-US23089.
PR 29-NOV-1999; 99WO-US28214.
PR 30-NOV-1999; 99WO-US28313.
PR 01-DEC-1999; 99WO-US28301.
PR 02-DEC-1999; 99WO-US28564.
PR 02-DEC-1999; 99WO-US28565.
PR 16-DEC-1999; 99WO-US30095.
PR 20-DEC-1999; 99WO-US30911.
PR 20-DEC-1999; 99WO-US30999.
PR 05-JAN-2000; 2000WO-US00219.
PR 11-FEB-2000; 2000WO-US03565.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US05004.
PR 02-MAR-2000; 2000WO-US05841.
PR 20-MAR-2000; 2000WO-US07377.
PR 30-MAR-2000; 2000WO-US08439.
PR 22-MAY-2000; 2000WO-US14042.
PR 02-JUN-2000; 2000WO-US15264.
PR 28-JUL-2000; 2000WO-US20710.
PR 24-AUG-2000; 2000WO-US23328.
PR 17-SEP-1997; 97US-059113P.
PR 17-SEP-1997; 97US-059115P.
PR 17-SEP-1997; 97US-059117P.
PR 17-SEP-1997; 97US-059119P.
PR 17-SEP-1997; 97US-059121P.
PR 17-SEP-1997; 97US-059122P.
PR 17-SEP-1997; 97US-059184P.
PR 18-SEP-1997; 97US-059263P.
PR 18-SEP-1997; 97US-059266P.
PR 15-OCT-1997; 97US-062125P.
PR 17-OCT-1997; 97US-062285P.
PR 17-OCT-1997; 97US-062287P.
PR 21-OCT-1997; 97US-063486P.
PR 24-OCT-1997; 97US-062814P.
PR 24-OCT-1997; 97US-062816P.
PR 24-OCT-1997; 97US-063045P.
PR 24-OCT-1997; 97US-063120P.
PR 24-OCT-1997; 97US-063121P.
PR 24-OCT-1997; 97US-063127P.
PR 24-OCT-1997; 97US-063128P.
PR 27-OCT-1997; 97US-063327P.
PR 27-OCT-1997; 97US-063329P.
PR 28-OCT-1997; 97US-063541P.
PR 28-OCT-1997; 97US-063542P.
PR 28-OCT-1997; 97US-063544P.
PR 28-OCT-1997; 97US-063549P.
PR 28-OCT-1997; 97US-063550P.
PR 28-OCT-1997; 97US-063564P.

PR 29-OCT-1997; 97US-063435P.
PR 29-OCT-1997; 97US-063704P.
PR 29-OCT-1997; 97US-063732P.
PR 29-OCT-1997; 97US-063734P.
PR 29-OCT-1997; 97US-063735P.
PR 29-OCT-1997; 97US-063738P.
PR 29-OCT-1997; 97US-064215P.
PR 31-OCT-1997; 97US-063870P.
PR 31-OCT-1997; 97US-064103P.
PR 03-NOV-1997; 97US-064248P.
PR 07-NOV-1997; 97US-064809P.
PR 12-NOV-1997; 97US-065186P.
PR 17-NOV-1997; 97US-065846P.
PR 18-NOV-1997; 97US-065693P.
PR 21-NOV-1997; 97US-066120P.
PR 21-NOV-1997; 97US-066364P.
PR 24-NOV-1997; 97US-066453P.
PR 24-NOV-1997; 97US-066466P.
PR 24-NOV-1997; 97US-066511P.
PR 24-NOV-1997; 97US-066770P.
PR 24-NOV-1997; 97US-066772P.
PR 18-SEP-2000; 2000US-0665350.
XX
PA (GETH) GENENTECH INC.
XX
PI Ashkenazi A, Botstein D, Desnoyers L, Eaton DL, Ferrara N;
PI Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME, Goddard A;
PI Godowski PJ, Grimaldi JC, Gurney AL, Hillan KJ, Kljavin IJ;
PI Mather JP, Pan J, Paoni NF, Roy MA, Stewart TA, Tumas D;
PI Williams PM, Wood WI;
XX
DR WPI; 2003-329602/31.
DR N-PSDB; ACA60112.
XX
PT New transmembrane polypeptides and nucleic acids encoding the
PT polypeptides, useful in gene therapy, in chromosome identification, as
PT chromosome markers, in generating probes and in tissue typing -
XX
PS Claim 12; Fig 50; 484pp; English.
XX
CC The invention relates to an isolated nucleic acid with at least 80%
CC nucleic acid sequence identity to a nucleotide sequence encoding one of
CC 61 secreted/transmembrane polypeptides, or PRO polypeptides or encoding a
CC PRO protein extracellular domain. Also included are a vector comprising
CC the PRO nucleic acid, a host cell comprising the vector, producing a PRO
CC polypeptide (by culturing the host cell for the expression of the PRO
CC polypeptide, and recovering the PRO polypeptide from the cell culture),
CC an isolated PRO polypeptide (having at least 80% sequence identity
CC to: (a) an amino acid sequence selected from the 61 PRO proteins;
CC (b) an amino acid sequence encoded by a nucleic acid molecule deposited
CC with an ATCC number (detailed in the specification); or (c) an
CC extracellular domain of a PRO polypeptide or to a PRO polypeptide lacking
CC its associated signal peptide), a chimaeric molecule comprising a PRO
CC polypeptide of fused to a heterologous amino acid sequence, an anti-PRO
CC antibody, detecting a PRO245 or PRO1868 in a sample suspected of
CC containing the polypeptide, linking a bioactive molecule to a cell
CC expressing a PRO245 or PRO1868 and modulating at least one biological
CC activity of a cell expressing a PRO245 or PRO1868. Nucleic acids which
CC encode PRO can be used to generate either transgenic animals or knock-out
CC animals which may be used in the development and screening of
CC therapeutically useful reagents. The nucleic acids may also be used in
CC gene therapy, in chromosome identification, as chromosome markers, or in
CC generating probes. The PRO polypeptides are useful as molecular markers
CC for protein electrophoresis, and the isolated nucleic acids may be used
CC for recombinantly expressing those markers. The PRO polypeptides and
CC nucleic acids may also be used in tissue typing. Anti-PRO antibodies
CC are useful in diagnostic assays for PRO, and in affinity purification
CC of PRO from recombinant cell culture or natural sources. The
XX present sequence represents a PRO protein.

SQ Sequence 382 AA;

Query Match

98.7%; Score 1101; DB 24; Length 382;

Best Local Similarity 96.3%; Pred. No. 9.6e-104;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 ATGRLLS-----GQVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
|||
Db 22 ATGRLLSASDLDLRGGQPVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81
|||
QY 53 ESEDEQKLIKFIENLLPSDGDGFWIGLRRRREKQSNSTACQDLYAWTDGSIQFRNYYVD 112
|||
Db 82 ESEDEQKLIKFIENLLPSDGDGFWIGLRRRREKQSNSTACQDLYAWTDGSIQFRNYYVD 141
|||
QY 113 EPSCGSEVCVVMYHQPSAPAGIGGYPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGE 172
|||
Db 142 EPSCGSEVCVVMYHQPSAPAGIGGYPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGE 201
|||
QY 173 ETELTTPVLPEETQEEADAKKTFKESREAAALNLAY 206
|||
Db 202 ETELTTPVLPEETQEEADAKKTFKESREAAALNLAY 235 ,
|||

RESULT 14
ABU65578
ID ABU65578 standard; Protein; 382 AA.
XX
AC ABU65578;
XX
DT 19-MAY-2003 (first entry)
XX
DE Human secreted/transmembrane protein, SEQ ID 20.
XX
KW Human; PRO; secreted protein; transmembrane protein;
KW cytosolic; antiarthritic; osteopathic; adrenal tumour; lung tumour;
KW colon tumour; breast tumour; prostate tumour; rectal tumour;
KW cervical tumour; liver tumour; TNF-alpha release; arthritis;
KW tumour necrosis factor alpha; chondrocyte cell; bone disorder;
KW cartilage disorder; sports injury.
XX
OS Homo sapiens.
XX
PN US2003036156-A1.
XY 20-FEB-2003.
PL
XX
PF 02-JUL-2002; 2002US-0188767.
XX 16-SEP-1998; 98WO-US19330.
PR 07-OCT-1998; 98WO-US21141.
PR 01-DEC-1998; 98WO-US25108.
PR 08-MAR-1999; 99WO-US05028.
PR 14-MAY-1999; 99WO-US10733.
PR 02-JUN-1999; 99WO-US12252.
PR 01-SEP-1999; 99WO-US20111.
PR 15-SEP-1999; 99WO-US21090.
PR 01-DEC-1999; 99WO-US28301.
PR 02-DEC-1999; 99WO-US28551.
PR 30-DEC-1999; 99WO-US31274.
PR 05-JAN-2000; 2000WO-US00219.
PR 18-FEB-2000; 2000WO-US04341.
PR 18-FEB-2000; 2000WO-US04342.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US05004.
PR 01-MAR-2000; 2000WO-US05601.
PR 02-MAR-2000; 2000WO-US05841.
PR 15-MAR-2000; 2000WO-US06884.
PR 30-MAR-2000; 2000WO-US08439.
PR 17-MAY-2000; 2000WO-US13705.
PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.
PR 02-JUN-2000; 2000WO-US15264.
PR 28-JUL-2000; 2000WO-US20710.
PR 24-AUG-2000; 2000WO-US23328.
PR 08-NOV-2000; 2000WO-US30952.
PR 01-DEC-2000; 2000WO-US32678.

PR 20-DEC-2000; 2000WO-US34956.
PR 28-FEB-2001; 2001WO-US06520.
PR 01-JUN-2001; 2001WO-US17800.
PR 20-JUN-2001; 2001WO-US19692.
PR 29-JUN-2001; 2001WO-US21066.
PR 09-JUL-2001; 2001WO-US21735.
PR 29-AUG-2001; 2001WO-US27099.
PR 18-SEP-1997; 97US-059263P.
PR 18-SEP-1997; 97US-059266P.
PR 17-OCT-1997; 97US-062250P.
PR 21-OCT-1997; 97US-063486P.
PR 24-OCT-1997; 97US-063120P.
PR 24-OCT-1997; 97US-063121P.
PR 28-OCT-1997; 97US-063540P.
PR 28-OCT-1997; 97US-063541P.
PR 28-OCT-1997; 97US-063544P.
PR 28-OCT-1997; 97US-063564P.
PR 29-OCT-1997; 97US-063734P.
PR 31-OCT-1997; 97US-063870P.
PR 31-OCT-1997; 97US-064103P.
PR 13-NOV-1997; 97US-065311P.
PR 21-NOV-1997; 97US-066120P.
PR 24-NOV-1997; 97US-066466P.
PR 11-DEC-1997; 97US-069335P.
PR 12-DEC-1997; 97US-069425P.
PR 17-DEC-1997; 97US-069870P.
PR 18-DEC-1997; 97US-068017P.
PR 10-MAR-1998; 98US-077450P.
PR 11-MAR-1998; 98US-077632P.
PR 11-MAR-1998; 98US-077649P.
PR 20-MAR-1998; 98US-078886P.
PR 20-MAR-1998; 98US-078939P.
PR 27-MAR-1998; 98US-079664P.
PR 27-MAR-1998; 98US-079786P.
PR 31-MAR-1998; 98US-080107P.
PR 31-MAR-1998; 98US-080194P.
PR 01-APR-1998; 98US-080327P.
PR 01-APR-1998; 98US-080333P.
PR 08-APR-1998; 98US-081049P.
PR 08-APR-1998; 98US-081070P.
PR 09-APR-1998; 98US-081195P.
PR 15-APR-1998; 98US-081838P.
PR 21-APR-1998; 98US-082568P.
PR 21-APR-1998; 98US-082569P.
PR 22-APR-1998; 98US-082704P.
PR 22-APR-1998; 98US-082797P.
PR 28-APR-1998; 98US-083322P.
PR 29-APR-1998; 98US-083495P.
PR 29-APR-1998; 98US-083496P.
PR 29-APR-1998; 98US-083499P.
PR 29-APR-1998; 98US-083559P.
PR 05-MAY-1998; 98US-084366P.
PR 06-MAY-1998; 98US-084414P.
PR 07-MAY-1998; 98US-084639P.
PR 07-MAY-1998; 98US-084640P.
PR 07-MAY-1998; 98US-084643P.
PR 15-MAY-1998; 98US-085579P.
PR 15-MAY-1998; 98US-085580P.
PR 15-MAY-1998; 98US-085582P.
PR 15-MAY-1998; 98US-085700P.
PR 18-MAY-1998; 98US-086023P.
PR 22-MAY-1998; 98US-086392P.
PR 22-MAY-1998; 98US-086486P.
PR 28-MAY-1998; 98US-087098P.
PR 28-MAY-1998; 98US-087208P.
PR 02-JUN-1998; 98US-087609P.
PR 02-JUN-1998; 98US-087759P.
PR 03-JUN-1998; 98US-087827P.
PR 04-JUN-1998; 98US-088025P.
PR 04-JUN-1998; 98US-088028P.
PR 04-JUN-1998; 98US-088029P.
PR 04-JUN-1998; 98US-088033P.

PR 04-JUN-1998; 98US-088326P.
PR 05-JUN-1998; 98US-088167P.
PR 05-JUN-1998; 98US-088202P.
PR 05-JUN-1998; 98US-088212P.
PR 05-JUN-1998; 98US-088217P.
PR 09-JUN-1998; 98US-088655P.
PR 10-JUN-1998; 98US-088722P.
PR 10-JUN-1998; 98US-088738P.
PR 10-JUN-1998; 98US-088740P.
PR 10-JUN-1998; 98US-088811P.
PR 10-JUN-1998; 98US-088824P.
PR 10-JUN-1998; 98US-088825P.
PR 10-JUN-1998; 98US-088826P.
PR 11-JUN-1998; 98US-088861P.
PR 11-JUN-1998; 98US-088863P.
PR 11-JUN-1998; 98US-088876P.
PR 12-JUN-1998; 98US-089090P.
PR 12-JUN-1998; 98US-089105P.
PR 16-JUN-1998; 98US-089512P.
PR 16-JUN-1998; 98US-089514P.
PR 17-JUN-1998; 98US-089538P.
PR 17-JUN-1998; 98US-089598P.
PR 17-JUN-1998; 98US-089653P.
PR 18-JUN-1998; 98US-089908P.
PR 19-JUN-1998; 98US-089952P.
PR 22-JUN-1998; 98US-090246P.
PR 22-JUN-1998; 98US-090252P.
PR 22-JUN-1998; 98US-090254P.
PR 24-JUN-1998; 98US-090429P.
PR 24-JUN-1998; 98US-090435P.
PR 24-JUN-1998; 98US-090444P.
PR 24-JUN-1998; 98US-090461P.
PR 24-JUN-1998; 98US-090535P.
PR 24-JUN-1998; 98US-090540P.
PR 25-JUN-1998; 98US-090676P.
PR 25-JUN-1998; 98US-090678P.
PR 25-JUN-1998; 98US-090688P.
PR 25-JUN-1998; 98US-090690P.
PR 25-JUN-1998; 98US-090694P.
PR 25-JUN-1998; 98US-090695P.
PR 25-JUN-1998; 98US-090696P.
PR 26-JUN-1998; 98US-090862P.
PR 26-JUN-1998; 98US-090863P.
PR 26-JUN-1998; 98US-091010P.
PR 01-JUL-1998; 98US-091359P.
PR 01-JUL-1998; 98US-091544P.
PR 02-JUL-1998; 98US-091478P.
PR 02-JUL-1998; 98US-091486P.
PR 02-JUL-1998; 98US-091626P.
PR 02-JUL-1998; 98US-091628P.
PR 02-JUL-1998; 98US-091632P.
PR 24-JUL-1998; 98US-094006P.
PR 04-AUG-1998; 98US-095282P.
PR 10-AUG-1998; 98US-095998P.
PR 10-AUG-1998; 98US-096012P.
PR 17-AUG-1998; 98US-096757P.
PR 17-AUG-1998; 98US-096766P.
PR 17-AUG-1998; 98US-096867P.
PR 17-AUG-1998; 98US-096891P.
PR 17-AUG-1998; 98US-096897P.
PR 18-AUG-1998; 98US-096949P.
PR 18-AUG-1998; 98US-096959P.
PR 18-AUG-1998; 98US-097022P.
PR 26-AUG-1998; 98US-097952P.
PR 26-AUG-1998; 98US-097954P.
PR 26-AUG-1998; 98US-097955P.
PR 26-AUG-1998; 98US-097971P.
PR 26-AUG-1998; 98US-097974P.
PR 26-AUG-1998; 98US-098014P.
PR 01-SEP-1998; 98US-098716P.
PR 01-SEP-1998; 98US-098723P.
PR 02-SEP-1998; 98US-098803P.
PR 02-SEP-1998; 98US-098821P.

PR 04-JUN-1998; 98US-088028P.
PR 04-JUN-1998; 98US-088029P.
PR 04-JUN-1998; 98US-088033P.
PR 04-JUN-1998; 98US-088326P.
PR 05-JUN-1998; 98US-088167P.
PR 05-JUN-1998; 98US-088202P.
PR 05-JUN-1998; 98US-088212P.
PR 05-JUN-1998; 98US-088217P.
PR 09-JUN-1998; 98US-088655P.
PR 10-JUN-1998; 98US-088722P.
PR 10-JUN-1998; 98US-088738P.
PR 10-JUN-1998; 98US-088740P.
PR 10-JUN-1998; 98US-088811P.
PR 10-JUN-1998; 98US-088824P.
PR 10-JUN-1998; 98US-088825P.
PR 10-JUN-1998; 98US-088826P.
PR 11-JUN-1998; 98US-088861P.
PR 11-JUN-1998; 98US-088863P.
PR 11-JUN-1998; 98US-088876P.
PR 12-JUN-1998; 98US-089090P.
PR 12-JUN-1998; 98US-089105P.
PR 16-JUN-1998; 98US-089512P.
PR 16-JUN-1998; 98US-089514P.
PR 17-JUN-1998; 98US-089538P.
PR 17-JUN-1998; 98US-089598P.
PR 17-JUN-1998; 98US-089653P.
PR 18-JUN-1998; 98US-089908P.
PR 19-JUN-1998; 98US-089952P.
PR 22-JUN-1998; 98US-090246P.
PR 22-JUN-1998; 98US-090252P.
PR 22-JUN-1998; 98US-090254P.
PR 24-JUN-1998; 98US-090429P.
PR 24-JUN-1998; 98US-090435P.
PR 24-JUN-1998; 98US-090444P.
PR 24-JUN-1998; 98US-090461P.
PR 24-JUN-1998; 98US-090535P.
PR 24-JUN-1998; 98US-090540P.
PR 25-JUN-1998; 98US-090678P.
PR 25-JUN-1998; 98US-090688P.
PR 25-JUN-1998; 98US-090690P.
PR 25-JUN-1998; 98US-090694P.
PR 25-JUN-1998; 98US-090695P.
PR 25-JUN-1998; 98US-090696P.
PR 26-JUN-1998; 98US-090862P.
PR 26-JUN-1998; 98US-090863P.
PR 26-JUN-1998; 98US-091010P.
PR 01-JUL-1998; 98US-091359P.
PR 01-JUL-1998; 98US-091544P.
PR 02-JUL-1998; 98US-091478P.
PR 02-JUL-1998; 98US-091486P.
PR 02-JUL-1998; 98US-091626P.
PR 02-JUL-1998; 98US-091628P.
PR 02-JUL-1998; 98US-091632P.
PR 24-JUL-1998; 98US-094006P.
PR 04-AUG-1998; 98US-095282P.
PR 10-AUG-1998; 98US-095998P.
PR 17-AUG-1998; 98US-096012P.
PR 17-AUG-1998; 98US-096757P.
PR 17-AUG-1998; 98US-096766P.
PR 17-AUG-1998; 98US-096867P.
PR 17-AUG-1998; 98US-096891P.
PR 17-AUG-1998; 98US-096897P.
PR 18-AUG-1998; 98US-096949P.
PR 18-AUG-1998; 98US-096959P.
PR 18-AUG-1998; 98US-097022P.
PR 26-AUG-1998; 98US-097952P.
PR 26-AUG-1998; 98US-097954P.
PR 26-AUG-1998; 98US-097955P.
PR 26-AUG-1998; 98US-097971P.
PR 26-AUG-1998; 98US-097974P.
PR 26-AUG-1998; 98US-098014P.
PR 01-SEP-1998; 98US-098716P.

PR 01-SEP-1998; 98US-098723P.
PR 02-SEP-1998; 98US-098803P.
PR 02-SEP-1998; 98US-098821P.
PR 02-SEP-1998; 98US-098843P.
PR 09-SEP-1998; 98US-099602P.
PR 10-SEP-1998; 98US-099741P.
PR 10-SEP-1998; 98US-099754P.
PR 10-SEP-1998; 98US-099763P.

Query Match 98.7%; Score 1101; DB 24; Length 382;
Best Local Similarity 96.3%; Pred. No. 9.6e-104;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 ATGRLLS-----GQPVCRGGTORPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
|||
Db 22 ATGRLLSASDLRLGGQPVCRGGTORPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81
|||
QY 53 ESEDEQKLIKFIENLLPSDGDWIGLRRRREEKQSNSTACQDLYAWTDGSIQFRNMYVD 112
|||
Db 82 ESEDEQKLIKFIENLLPSDGDWIGLRRRREEKQSNSTACQDLYAWTDGSIQFRNMYVD 141
|||
QY 113 EPSCGSEVCVMYHQPSAPAGIGGPFQWQNDRCNMKNNFICKYSDEKPAVPSREAEGE 172
|||
Db 142 EPSCGSEVCVMYHQPSAPAGIGGPFQWQNDRCNMKNNFICKYSDEKPAVPSREAEGE 201
|||
QY 173 ETELTPVLPEETQEEDAKKTFKESREAAALNLAY 206
|||
Db 202 ETELTPVLPEETQEEDAKKTFKESREAAALNLAY 235
|||

Search completed: December 22, 2003, 16:10:09
Job time : 29.1819 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 22, 2003, 16:05:39 ; Search time 24.0017 Seconds
(without alignments)
2214.797 Million cell updates/sec

Title: US-09-887-855-2_COPY_22_227
Perfect score: 1115
Sequence: 1 ATGRLLSGQPVCRCGGTQPC.....EEDAKKTFKESREAAALNLAY 206

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvrius:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1115	100.0	374	4	Q8TAY8	Q8tay8 homo sapien
2	1115	100.0	374	4	Q96NF3	Q96nf3 homo sapien
3	1111	99.6	374	4	Q96NC5	Q96nc5 homo sapien
4	945.5	84.8	374	11	Q9Z209	Q9z209 cricetus
5	803.5	72.1	211	11	Q8C351	Q8c351 mus musculus
6	580	52.0	246	11	Q8BVI7	Q8bvi7 mus musculus
7	580	52.0	292	11	Q8BVU2	Q8bv12 mus musculus
8	185	16.6	1290	13	Q9W6E1	Q9w6e1 gallus gall
9	178.5	16.0	1456	11	Q61830	Q61830 mus musculus
10	177.5	15.9	1348	5	Q25199	Q25199 hydra atten
11	176.5	15.8	1479	4	Q9Y5P9	Q9y5p9 homo sapien
12	176.5	15.8	1479	4	Q9UBG0	Q9ubg0 homo sapien
13	170.5	15.3	1479	11	Q64449	Q64449 mus musculus
14	169	15.2	217	11	Q8C4F8	Q8c4f8 mus musculus
15	165.5	14.8	742	11	Q8K4Q8	Q8k4q8 mus musculus
16	165.5	14.8	742	11	Q8C979	Q8c979 mus musculus

17	161.5	14.5	134	5	Q9XYX3	Q9xyx3 hydra magni
18	159.5	14.3	742	11	Q8VIF6	Q8vif6 mus musculu
19	158.5	14.2	719	6	O62623	O62623 bos taurus
20	158	14.2	142	11	Q8CJ86	Q8cj86 mus musculu
21	158	14.2	142	11	Q8BHK7	Q8bhk7 mus musculu
22	158	14.2	295	11	Q91ZM4	Q91zm4 mus musculu
23	158	14.2	311	11	Q9D8V4	Q9d8v4 mus musculu
24	158	14.2	325	11	Q91ZX0	Q91zx0 mus musculu
25	157	14.1	158	13	Q90WI7	Q90wi7 bungarus fa
26	156	14.0	323	11	Q8CJ91	Q8cj91 mus musculu
27	156	14.0	339	6	Q95244	Q95244 sus scrofa
28	155	13.9	293	11	Q8BGZ0	Q8bgz0 mus musculu
29	155	13.9	323	11	Q8CJ94	Q8cj94 mus musculu
30	155	13.9	323	11	Q8CJ93	Q8cj93 mus musculu
31	155	13.9	323	11	Q8CJ88	Q8cj88 mus musculu
32	155	13.9	1152	13	Q90WM2	Q90wm2 xenopus lae
33	154	13.8	322	11	Q8CJ89	Q8cj89 mus musculu
34	154	13.8	323	11	Q8CJ92	Q8cj92 mus musculu
35	152.5	13.7	485	6	Q95LG3	Q95lg3 odocoileus
36	152	13.6	158	13	Q90WI6	Q90wi6 bungarus mu
37	151.5	13.6	652	4	Q8IXK1	Q8ixk1 homo sapien
38	150.5	13.5	399	6	Q8HY12	Q8hy12 hylobates l
39	150	13.5	197	6	Q28008	Q28008 bos taurus
40	149.5	13.4	158	11	Q8JZX6	Q8jzx6 mus musculu
41	149.5	13.4	459	5	Q22136	Q22136 caenorhabdi
42	149	13.4	158	13	Q90WI8	Q90wi8 bungarus fa
43	148.5	13.3	163	13	Q8AXR8	Q8axr8 anguilla ja
44	148.5	13.3	195	5	Q27340	Q27340 megabalanus
45	148.5	13.3	330	4	Q9NT67	Q9nt67 homo sapien

ALIGNMENTS

RESULT 1

Q8TAY8 PRELIMINARY; PRT; 374 AA.
AC Q8TAY8;
DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
D 01-OCT-2002 (Tremblrel. 22, Last annotation update)
DE Similar to unnamed protein product.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC025407; AAH25407.1; -.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00059; lectin_c; 1.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS50041; C TYPE LECTIN 2; 1.
SQ SEQUENCE 374 AA; 42312 MW; FC214E6BC9E578D9 CRC64;

Query Match 100.0%; Score 1115; DB 4; Length 374;
Best Local Similarity 100.0%; Pred. NO. 4.5e-99;
Matches 206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGRLLSGQPVCRCGGTQPCYKVIYFHDTSRLNFEAKEACRRDGGQLVSISEDEQKL 60
Db	22	ATGRLLSGQPVCRCGGTQPCYKVIYFHDTSRLNFEAKEACRRDGGQLVSISEDEQKL 81
QY	61	IEKFIENLLPSDGFWIGLRRREKQSNSTACQDLYAWTDGSIQFRNWWYDEPSCGSEV 120
Db	82	IEKFIENLLPSDGFWIGLRRREKQSNSTACQDLYAWTDGSIQFRNWWYDEPSCGSEV 141
QY	121	CVVMYHQPSAPAGTGGPYMFQWDDRCNMKNFNICKYSDEKPAVPSREAGEETEELTPV 180
Db	142	CVVMYHQPSAPAGTGGPYMFQWDDRCNMKNFNICKYSDEKPAVPSREAGEETEELTPV 201

QY 63 KFIENLLPSDGFWIGLRRRREKQSNSTACODLYAWTDGSIQFRNWWYVDEPSCGSEVCV 122
Db 84 KFIENLLASDGFWIGLRRLEVKQVNTACQDLYAWTDGTSQFRNWWYVDEPSCGSEVCV 143
QY 123 VMYHQPSAPAGIGGYPYMFQWDDRCNMKNFNICKYSDEKPA-VPSREAEGETELTTPVL 181
Db 144 VMYHQPSAPPGIGGYPYMFQWDDRCNMKNFNICKYADEKPSSTPSIRPGSEATEPPTPVL 203
QY 182 PEETOEDAKKTFKESREAAALNAY 206
Db 204 PEETOKEDTKTFKESREAAALNAY 228

RESULT 5
Q8C351 PRELIMINARY; PRT; 211 AA.
ID Q8C351
AC Q8C351;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE LAYILIN homolog (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Lung;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL; AK086930; BAC39765.1; --
FT NON TER 211
SQ SEQUENCE 211 AA; 23697 MW; AD9870B5957DD5AE CRC64;

Query Match 72.1%; Score 803.5; DB 11; Length 211;
Best Local Similarity 79.8%; Pred. No. 2.2e-69;
Matches 150; Conservative 9; Mismatches 20; Indels 9; Gaps 2;
QY 3 GRLLS-----GQVCRGGTQRPCYKVIYFHDTSRRLNFEAEKACRRDGGQLVSIES 54
Db 24 GRLLSASDLDPGGQLVCRGGTTRRPCYKVIYFHDFAQRLNFEAEKTCMEDGGQLVSIE 83
QY 55 EDEQKLIKFIENLLPSDGFWIGLRRRREKQSNSTACQDLYAWTDGSIQFRNWWYVDEP 114
Db 84 EDEQRLIEFIENLLASDGFWIGLRLKEEKQSNNTACQDLYAWTDGTSQFRNWWYVDEP 143
QY 115 SCGSEVCVMYHQPSAPAGIGGYPYMFQWDDRCNMKNFNICKYSDEKPA-VPSREAESEE 173
Db 144 SCGSEVCVMYHQPSAPPGIGGYPYMFQWDDRCNMKNFNICKYHDDKPSSTPSIWPGEA 203
QY 174 TELTTPVL 181
Db 204 TEPATPLL 211

RESULT 6
Q8BMI7 PRELIMINARY; PRT; 246 AA.
ID Q8BMI7
AC Q8BMI7;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE C-type-lectin protein MT75 homolog.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Forelimb;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL; AK031063; BAC27234.1; --
SQ SEQUENCE 246 AA; 27394 MW; E872660C58267752 CRC64;
Query Match 52.0%; Score 580; DB 11; Length 246;
Best Local Similarity 57.9%; Pred. No. 8.9e-48;
Matches 110; Conservative 26; Mismatches 40; Indels 14; Gaps 5;
QY 4 RLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEAEKACRRDGGQLVSISEDEQKLIK 63
Db 23 RVVSGQKVCADVKHPCYKMAFYHELSSRSVSFQEARLACESEGGVLLSLENAEQKLIES 82
QY 64 FIENLLP-----SDGDFWIGLRRRREKQSNSTACQDLYAWTDGSIQFRNWWYVDEPSCGS 118
Db 83 MLQNLTKPGTGISDGFWIGLRLSGDGT-SGACPDLYQWSDGSSSQFRNWWYTDPEPSCGS 141
QY 119 EVCVVMYHQPSAPAGIGGYPYMFQWDDRCNMKNFNICKYSDE-KPAVPSREAEGETELT 177
Db 142 EKCVMYHQPTANPGLGGPYLYQWDDRCNMKNHYICKYEPIHPTPEA-----EKPYL 196
QY 178 TPVLPETQE 187
Db 197 NQ--PEETHE 204

RESULT 7
Q8BVU2 PRELIMINARY; PRT; 292 AA.
ID Q8BVU2
AC Q8BVU2;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE C-type lectin protein MT75 homolog.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL; AK076523; BAC36378.1; --
SQ SEQUENCE 292 AA; 32502 MW; 73E631C0714D54E2 CRC64;

Query Match 52.0%; Score 580; DB 11; Length 292;
Best Local Similarity 57.9%; Pred. No. 1.1e-47;
Matches 110; Conservative 26; Mismatches 40; Indels 14; Gaps 5;
QY 4 RLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEAEKACRRDGGQLVSISEDEQKLIK 63
Db 23 RVVSGQKVCADVKHPCYKMAFYHELSSRSVSFQEARLACESEGGVLLSLENAEQKLIES 82
QY 64 FIENLLP-----SDGDFWIGLRRRREKQSNSTACQDLYAWTDGSIQFRNWWYVDEPSCGS 118
Db 83 MLQNLTKPGTGISDGFWIGLRLSGDGT-SGACPDLYQWSDGSSSQFRNWWYTDPEPSCGS 141
QY 119 EVCVVMYHQPSAPAGIGGYPYMFQWDDRCNMKNFNICKYSDE-KPAVPSREAEGETELT 177
Db 142 EKCVMYHQPTANPGLGGPYLYQWDDRCNMKNHYICKYEPIHPTPEA-----EKPYL 196
QY 178 TPVLPETQE 187


```
Db      197 NQ--PEETHE 204
|||||
RESULT 8
Q9W6E1
ID Q9W6E1 PRELIMINARY; PRT; 1290 AA.
AC Q9W6E1;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Neurocan core protein.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A. PubMed=10851024;
RX MEDLINE=20309833; Hoffmann S., Balsamo J., Lilien J.;
RA Li H., Leung T.C., Hoffman S., Balsamo J., Lilien J.;
RT "Coordinate Regulation of Cadherin and Integrin Function by the
RT Chondroitin Sulfate Proteoglycan Neurocan.";
RL J. Cell Biol. 149:1275-1288(2000).
DR EMBL; AF116856; AAD24546.2; -.
DR HSSP; P08709; 1BF9.
DR InterPro; IPR00152; Asx_hydroxyl.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR001438; EGF_II.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR000538; Link.
DR InterPro; IPR000436; Sushi_SCR_CCP.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF00059; lectin_c; 1.
DR Pfam; PF00084; sushi; 1.
DR Pfam; PF00084; sushi; 1.
DR Pfam; PF00193; xlink; 2.
DR PRINTS; PR00010; EGFBL00D.
DR PRINTS; PR01265; LINKMODULE.
DR ProDom; PD000918; Link; 2.
DR SMART; SM00032; CCP; 1.
DR SMART; SM00034; CLECT; 1.
DR SMART; SM00179; EGF_CA; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00445; LINK; 2.
DR PROSITE; PS00010; ASX_HYDROXYL; 1.
DR PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR PROSITE; PS00022; EGF_1_2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01187; EGF_CA; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR PROSITE; PS01241; LINK; 2.
KW EGF-like domain.
SQ SEQUENCE 1290 AA; 138877 MW; 182BD86D0E40BE78 CRC64;

Query Match 16.6%; Score 185; DB 13; Length 1290;
Best Local Similarity 34.7%; Pred. No. 7e-09;
Matches 50; Conservative 18; Mismatches 42; Indels 34; Gaps 8;

QY 17 QRPCYKVIYFHDTSRRLNFEAKEACRRDGGQLVSISEDEQKLIKFIENLLPSDGDWF 76
Db 1064 QGHCYR--YF---SRRRSWEDAERDCRRRAGHLTSHSQEEHGFINSF-----GHENTW 1112
QY 77 IGLRRREKQSNSTACQDLVYAWTDGSIQFRNWWYVDEPS---CGSEVCVVMY-HQPSAPA 132
Db 1113 IGLNDRIVEQD-----FQWTDNTGLQYENWRENQPDNFFAGGEDCVVLVSHE----- 1159
```

```
QY 133 GIGGPFYMFQWDDRCNMKNKFNICK 156
Db 1160 -IG-----KWDVPCNYPICK 1177

RESULT 9
Q61830
ID Q61830 PRELIMINARY; PRT; 1456 AA.
AC Q61830;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Macrophage mannose receptor precursor.
GN MRC1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6;
RX MEDLINE=93043353; PubMed=1421407;
RA Harris N., Rits M., Chang G., Ezekowitz R.B.;
RT "Characterization of the murine macrophage mannose receptor.";
RL Blood 80:2363-2373(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6;
RA Super M.;
RL Submitted (APR-1992) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z11974; CAA78028.1; -.
DR HSSP; P22897; 1EGG.
DR MGD; MGI:97142; Mrc1.
DR InterPro; IPR002353; AntifreezeII.
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR000772; Ricin_B_lectin.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00059; lectin_c; 8.
DR Pfam; PF00652; Ricin_B_lectin; 2.
DR PRINTS; PR00356; ANTIFREEZEII.
DR PRINTS; PR00013; FNTYPEII.
DR ProDom; PD000995; FN_Type_II; 1.
DR SMART; SM00034; CLECT; 8.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00458; RICIN; 1.
DR PROSITE; PS00615; C_TYPE_LECTIN_1; 6.
DR PROSITE; PS50041; C_TYPE_LECTIN_2; 8.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS50231; RICIN_B_LECTIN; 1.
KW Receptor; Signal.
FT SIGNAL 1 18 POTENTIAL.
FT CHAIN 19 1456 MACROPHAGE MANNOSE RECEPTOR.
SQ SEQUENCE 1456 AA; 165065 MW; 4EBD3F1B8619A594 CRC64;

Query Match 16.0%; Score 178.5; DB 11; Length 1456;
Best Local Similarity 25.9%; Pred. No. 3.4e-08;
Matches 53; Conservative 33; Mismatches 66; Indels 53; Gaps 8;

QY 21 YKVIYFHDTSRRLNFEAKEACRRDGGQLVSISEDEQKLIKFIENLLPSDGDWF 80
Db 807 YKDYQYFYSKEKETMDNARRCKKNFGDLATIKSEKKFLWKYI-NKNGGQSPYFIGML 865
QY 81 RREEKQSNSTACQDLVYAWTDGSIQFRNWWYVDEPSCGS--EVCVMYHQPSAPAGIGPY 138
Db 866 ISMDKK-----FIWMDGSKVDFAVATGEPNFANDDENCVTMY-----TNSGF---- 908
QY 139 MFQWDDRCNMKNKFNICK---YSDEKPAVPSREAEGETELTTPVLPEETQE----- 187
Db 909 ---WNDINCGYPNPFICQRHNSINATAMP-----TPTTPGCKEGWHLYKNK 954
QY 188 -----EDAKTFKESREAAALNL 204
: ||:::|:|
```

```
Db 955 CFKIFGFANEKKSQDARQACKGL 979

RESULT 10
Q25199
ID Q25199 PRELIMINARY; PRT; 1348 AA.
AC Q25199;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Tyrosine kinase receptor.
OS Hydra attenuata (Hydra) (Hydra vulgaris).
OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Anthomedusae;
OC Hydridae; Hydra.
OX NCBI_TaxID=6087;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Irvine;
RX MEDLINE=20209407; PubMed=10744720;
RA Reidling J.C., Miller M.A., Steele R.E.;
RT "Sweet Tooth, a Novel Receptor Protein-tyrosine Kinase with C-type
Lectin-like Extracellular Domains.";
RL J. Biol. Chem. 275:10323-10330(2000).
DR EMBL; L22612; AAA29218.2; -.
DR HSSP; P11362; LFCK.
DR InterPro; IPR001304; LECTIN_C.
DR InterPro; IPR003990; Pancreatias ac.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00059; lectin_c; 4.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR01504; PNCREATITSAP.
DR ProDom; PD000001; Prot_kinase; 2.
DR SMART; SM00034; CLECT; 4.
DR SMART; SM00219; TyrKC; 1.
DR PROSITE; PS00615; C_TYPE_LLECTIN_1; 2.
DR PROSITE; PS00615; C_TYPE_LLECTIN_2; 4.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
KW ATP-binding; Kinase; Receptor; Transferase; Tyrosine-protein kinase.
SQ SEQUENCE 1348 AA; 156916 MW; 77D2122093227FEF CRC64;

Query Match 15.9%; Score 177.5; DB 5; Length 1348;
Best Local Similarity 28.0%; Pred. No. 3.9e-08;
Matches 47; Conservative 32; Mismatches 46; Indels 43; Gaps 8;

QY 6 LSGQPVC--RGGTQRPC-----YKVIYFHDTSRRLLNFEAEKACRRDGGQLVSISEDE 57
Db 412 LSHRFICKVKRATNEYCAEGWTSYRIYCYFIYSIEFDWFKSFSCQNIGNLLSIENQEE 471
QY 58 QKLIEKFIENLLPSDGD-FWIGLRR-----REEKQSNSTACQDLYAWTDGSIQFNWY 110
Db 472 ----NRFIENDLIKNDKYIWIGLKNKIWDYLNKKNR-----FEWSDNTYTQFNWI 518
QY 111 VDEP--SCGSEVCVVMYHQSPAPAGIGGYPYMFQWNDRCNMKNNFICK 156
Db 519 TNQPDNNNGIESCEVMNYN-----GWSDECKVLNGFICK 553

RESULT 11
Q9Y5P9
ID Q9Y5P9 PRELIMINARY; PRT; 1479 AA.
AC Q9Y5P9;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Endocytic receptor Endo180.
GN ENDO180.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;

QY 7 SQQPVCRRGTTQPCYKVIYFHDTSRRLLNFEAEKACRRDGGQLVSISEDEQKLIKIEFIE 66
Db 385 SQQPP-----QGHCVRL-----QAEKRSWQESKKACLRGGDLVSIHSMAELEFITKQIK 434
QY 67 NLLPSDGDWFWIGLRRRREEKQSNSTACQDLYAWTDGSIQFNWYVDEPS---CGSEVCVV 123
Db 435 QEVE---ELWIGL-----NDLKLQMFWSGSLVSTHWHHPFPNPNFRDSDLCVT 483
QY 124 MYHQSPAPAGIGGYPYMFQWNDRCNMKNNFICKYSDEKPAVPSREAE 171
Db 484 IW----GPEG-----RWNDSPCNQSLPSICKKAGQLSQGAEEEDHG 520

RESULT 12
Q9UBG0
ID Q9UBG0 PRELIMINARY; PRT; 1479 AA.
AC Q9UBG0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Urokinase receptor-associated protein UPARAP.
GN KIAA0709.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Behrendt N., Jensen O.N., Engelholm L.H., Mortz E., Mann M., Dano K.;
RT "A urokinase receptor-associated protein with specific collagen-
binding properties.";
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=98403880; PubMed=9734811;
RA Ishikawa K., Nagase T., Suyama M., Miyajima N., Tanaka A., Kotani H.,
RA Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. X.
The complete sequences of 100 new cDNA clones from brain which can
```

```
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20148849; PubMed=10683150;
RA Sheikh H., Yarwood H., Ashworth A., Isacke C.;
RT "Endo180, an endocytic recycling glycoprotein related to the
RT macrophage mannose receptor is expressed on fibroblasts, endothelial
RT cells and macrophages and functions as a lectin receptor.";
RL J. Cell Sci. 113:1021-1032(2000).
DR EMBL; AF134838; AAD30280.1; -.
DR HSSP; P02751; 2FN2.
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR000566; Lipocln_cytFABP.
DR InterPro; IPR000772; Ricin_B_lectin.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00059; lectin_c; 8.
DR PRINTS; PR00013; FNTYPEII.
DR ProDom; PD000395; FN_Type_II; 1.
DR SMART; SM00034; CLECT; 8.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00458; RICIN; 1.
DR PROSITE; PS00615; C_TYPE_LLECTIN_1; 3.
DR PROSITE; PS00041; C_TYPE_LLECTIN_2; 8.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00213; LIPOCALIN; 1.
DR PROSITE; PS50231; RICIN_B_LLECTIN; 1.
KW Receptor.
SQ SEQUENCE 1479 AA; 166669 MW; 9F4BAF355F036FCE CRC64;

Query Match 15.8%; Score 176.5; DB 4; Length 1479;
Best Local Similarity 30.4%; Pred. No. 5.4e-08;
Matches 51; Conservative 25; Mismatches 57; Indels 35; Gaps 7;
```

RT code for large proteins in vitro.";
RL DNA Res. 5:169-176(1998).
DR EMBL; AF107292; AAF14192.1; -.
DR EMBL; AB014609; BAA31684.1; -.
DR HSP; P02751; 2FN2.
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR000566; Lipocln_cytFABP.
DR InterPro; IPR000772; Ricin_B_lectin.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00049; lectin.c; 8.
DR PRINTS; PR00013; FNTYPEII.
DR ProDom; PD000995; FN_Type_II; 1.
DR SMART; SM00034; CLECT; 8.
DR SMART; SM00458; RICIN; 1.
DR PROSITE; PS00615; C_TYPE_LECTIN_1; 3.
DR PROSITE; PS00041; C_TYPE_LECTIN_2; 8.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00213; LIPOCALIN; 1.
DR PROSITE; PS50231; RICIN_B_LECTIN; 1.
KW Kinase; Receptor.
SQ SEQUENCE 1479 AA; 166654 MW; C7583EA78E2792D1 CRC64;

Query Match 15.8%; Score 176.5; DB 4; Length 1479;
Best Local Similarity 30.4%; Pred. No. 5.4e-08;
Matches 51; Conservative 25; Mismatches 57; Indels 35; Gaps 7;

QY 7 SQQVCRGGTQPCYKVIYFHDTSRRLNFEFEAKACRRDGGQLVSIIESEDEQKLIKIEFIE 66
Db 385 SWQPF-----QGHCYRL-----QA EKRSWQESKACLRGGDLVSIHMAELEFITKQIK 434

QY 67 NLLPSDGDGFWIGLRRREKQSNSTACQDLYAWTDGSIQFRNWWYVDEPS---CGSEVCVV 123
Db 435 QEVE---ELWIGL-----NDLKLQMNFEWSGSLVSFTHWHPPFPNFRDLSLEDCVT 483

QY 124 MYHQSPAPAGIGGPFYMFQWNDRCNMKNFICKYSDEKPAVPSREAEG 171
Db 484 IW----GPEG-----RWNDSPCNQSLPSICKKAGQLSQGAAEEDHG 520

RESULT 13
Q64449 ID Q64449 PRELIMINARY; PRT; 1479 AA.
AC Q64449;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Lectin lambda.
GN MRC2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96355501; PubMed=8702911;
RA Wu K., Yuan J., Lasky L.A.;
RT "Characterization of a novel member of the macrophage mannose receptor
type C lectin family.";
RL J. Biol. Chem. 271:21323-21330(1996).
DR EMBL; U56734; AAC52729.1; -.
DR HSP; P02751; 2FN2.
DR MGD; MGI:107818; Mrc2.
DR InterPro; IPR002353; AntifreezeII.
DR InterPro; IPR001128; Cytochrome P450.
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR000566; Lipocln_cytFABP.
DR InterPro; IPR000772; Ricin_B_lectin.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00059; lectin.c; 8.
DR PRINTS; PR00356; ANTIFREEZEII.

DR PRINTS; PR00013; FNTYPEII.
DR ProDom; PD000995; FN_Type_II; 1.
DR SMART; SM00034; CLECT; 8.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00458; RICIN; 1.
DR PROSITE; PS00086; CYTOCHROME P450; 1.
DR PROSITE; PS00615; C_TYPE_LECTIN_1; 3.
DR PROSITE; PS00041; C_TYPE_LECTIN_2; 8.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00213; LIPOCALIN; 1.
DR PROSITE; PS50231; RICIN_B_LECTIN; 1.
SQ SEQUENCE 1479 AA; 167112 MW; 62D456E10B9B48C1 CRC64;

Query Match 15.3%; Score 170.5; DB 11; Length 1479;
Best Local Similarity 31.4%; Pred. No. 2.1e-07;
Matches 48; Conservative 22; Mismatches 48; Indels 35; Gaps 7;

QY 7 SQQVCRGGTQPCYKVIYFHDTSRRLNFEFEAKACRRDGGQLVSIIESEDEQKLIKIEFIE 66
Db 384 SWQPF-----QGHCYRL-----QA EKRSWQESKACLRGGDLVSIHMAELEFITKQIK 433

QY 67 NLLPSDGDGFWIGLRRREKQSNSTACQDLYAWTDGSIQFRNWWYVDEPS---CGSEVCVV 123
Db 434 QEVE---ELWIGL-----NDLKLQMNFEWSGSLVSFTHWHPPFPNFRDLSLEDCVT 482

QY 124 MYHQSPAPAGIGGPFYMFQWNDRCNMKNFICK 156
Db 483 IW----GPEG-----RWNDSPCNQSLPSICK 504

RESULT 14
Q8C4F8 ID Q8C4F8 PRELIMINARY; PRT; 217 AA.
AC Q8C4F8;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Chondroitin sulfate proteoglycan 3 (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK082298; BAC38458.1; -.
FT NON TER 1
SQ SEQUENCE 217 AA; 25949 MW; 5F1A6A57B05B76D6 CRC64;

Query Match 15.2%; Score 169; DB 11; Length 217;
Best Local Similarity 31.1%; Pred. No. 2.9e-08;
Matches 42; Conservative 16; Mismatches 47; Indels 30; Gaps 5;

QY 25 YFHDTSRRLNFEFEAKACRRDGGQLVSIIESEDEQKLIKFIENLLPSDGDGFWIGLRRREE 84
Db 3 YF---AHRRAWEDAERDCRRRAGHLTSVHSPEEHKFINSF-----GHENSWIGLNDRTV 53

QY 85 KQSNSTACQDLYAWTDGSIQFRNWWYVDEPS---CGSEVCVVMYHQSPAPAGIGGPFYMFQ 141
Db 54 ERD-----FQWTDNTGLQYENWREKQPDNFFAGGDCVVMVAHESG-----R 95

QY 142 WNDRCNMKNFICK 156
Db 96 WNDVPCNYNLPYVCK 110

RESULT 15

Search completed: December 22, 2003, 16:13:30
Job time : 25.2017 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 22, 2003, 16:05:39 ; Search time 23.7687 Seconds
(without alignments)
2214.797 Million cell updates/sec

Title: US-09-887-855-2_COPY_24_227
Perfect score: 1106
Sequence: 1 GRLLSGQPVCRGGTQPCYK.....EEDAKKTFKESREAAALNLAY 204

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL 23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1106	100.0	374	4 Q8TAY8	Q8tay8 homo sapien
2	1106	100.0	374	4 Q96NF3	Q96nf3 homo sapien
3	1102	99.6	374	4 Q96NC5	Q96nc5 homo sapien
4	945.5	85.5	374	11 Q9Z209	Q9z209 cricetulus
5	803.5	72.6	211	11 Q8C351	Q8c351 mus musculus
6	580	52.4	246	11 Q8BMT7	Q8bmt7 mus musculus
7	580	52.4	292	11 Q8BVU2	Q8bv2 mus musculus
8	185	16.7	1290	13 Q9W6E1	Q9w6e1 gallus galli
9	178.5	16.1	1456	11 Q61830	Q61830 mus musculus
10	177.5	16.0	1348	5 Q25199	Q25199 hydra atten
11	176.5	16.0	1479	4 Q9Y5P9	Q9y5p9 homo sapien
12	176.5	16.0	1479	4 Q9UBG0	Q9ubg0 homo sapien
13	170.5	15.4	1479	11 Q64449	Q64449 mus musculus
14	169	15.3	217	11 Q8C4F8	Q8c4f8 mus musculus
15	165.5	15.0	742	11 Q8K4Q8	Q8k4q8 mus musculus
16	165.5	15.0	742	11 Q8C979	Q8c979 mus musculus

17	161.5	14.6	134	5 Q9XYX3	Q9xyx3 hydra magni
18	159.5	14.4	742	11 Q8VIF6	Q8vif6 mus musculus
19	158.5	14.3	719	6 O62623	O62623 bos taurus
20	158	14.3	142	11 Q8CJ86	Q8cj86 mus musculus
21	158	14.3	142	11 Q8BH7	Q8bhk7 mus musculus
22	158	14.3	295	11 Q91ZM4	Q91zw4 mus musculus
23	158	14.3	311	11 Q9D8V4	Q9d8v4 mus musculus
24	158	14.3	325	11 Q91ZX0	Q91zx0 mus musculus
25	157	14.2	158	13 Q90W17	Q90wi7 bungarus fa
26	156	14.1	323	11 Q8CJ91	Q8cj91 mus musculus
27	156	14.1	339	6 Q95244	Q95244 sus scrofa
28	155	14.0	293	11 Q8BGZ0	Q8bgz0 mus musculus
29	155	14.0	323	11 Q8CJ94	Q8cj94 mus musculus
30	155	14.0	323	11 Q8CJ93	Q8cj93 mus musculus
31	155	14.0	323	11 Q8CJ88	Q8cj88 mus musculus
32	155	14.0	1152	13 Q90WM2	Q90wm2 xenopus lae
33	154	13.9	322	11 Q8CJ89	Q8cj89 mus musculus
34	154	13.9	323	11 Q8CJ92	Q8cj92 mus musculus
35	152.5	13.8	485	6 Q95LG3	Q95lg3 odocoileus
36	152	13.7	158	13 Q90W16	Q90wi6 bungarus mu
37	151.5	13.7	652	4 Q8IXK1	Q8ixk1 homo sapien
38	150.5	13.6	399	6 Q8HY12	Q8hy12 hylobates l
39	150	13.6	197	6 Q28008	Q28008 bos taurus
40	149.5	13.5	158	11 Q8JZX6	Q8jzx6 mus musculus
41	149.5	13.5	459	5 Q22136	Q22136 caenorhabdi
42	149	13.5	158	13 Q90W18	Q90wi8 bungarus fa
43	148.5	13.4	163	13 Q8AXR8	Q8axr8 anguilla ja
44	148.5	13.4	195	5 Q27340	Q27340 megabalanus
45	148.5	13.4	330	4 Q9NT67	Q9nt67 homo sapien

ALIGNMENTS

RESULT 1

Q8TAY8 PRELIMINARY; PRT; 374 AA.
AC Q8TAY8;
DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)
DE Similar to unnamed protein product.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBDJ databases.
DR EMBL; BC025407; AAH25407.1; -
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00059; lectin_c; 1.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS50041; C_TYPE LECTIN 2; 1.
SQ SEQUENCE 374 AA; 42312 MW; FC214E6BC9E578D9 CRC64;

Query Match 100.0%; Score 1106; DB 4; Length 374;
Best Local Similarity 100.0%; Pred. NO. 2.8e-98;
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	GRLLSGQPVCRGGTQPCYKVIYFHDTSRRLLNFEEAKEACRRDGGQLVSISEDEQKLI	60
Db	24	GRLLSGQPVCRGGTQPCYKVIYFHDTSRRLLNFEEAKEACRRDGGQLVSISEDEQKLI	83
QY	61	KFIENLLPSDGFWIGLRRREEKQSNSTACQDLYAWTDGSIISQFRNWWYVDEPSCGSEVCV	120
Db	84	KFIENLLPSDGFWIGLRRREEKQSNSTACQDLYAWTDGSIISQFRNWWYVDEPSCGSEVCV	143
QY	121	VMYHQPAPAGIGGPPYMFQWNDRCNMKNFICKYDEKPAVPSREAEGETELTPVLP	180
Db	144	VMYHQPAPAGIGGPPYMFQWNDRCNMKNFICKYDEKPAVPSREAEGETELTPVLP	203


```
Db      197 NQ--PEETHE 204
|||||
RESULT 8
Q9W6E1
ID      Q9W6E1      PRELIMINARY;      PRT; 1290 AA.
AC      Q9W6E1;
DT      01-NOV-1999 (TremBLrel. 12, Created)
DT      01-NOV-1999 (TremBLrel. 12, Last sequence update)
DT      01-MAR-2003 (TremBLrel. 23, Last annotation update)
DE      Neurocan core protein.
OS      Gallus gallus (Chicken).
OC      Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC      Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC      Gallus.
OX      NCBI_TaxID=9031;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=20309833; PubMed=10851024;
RA      Li H., Leung T C., Hoffman S., Balsamo J., Lilien J.;
RT      "Coordinate Regulation of Cadherin and Integrin Function by the
RT      Chondroitin Sulfate Proteoglycan Neurocan.";
RL      J. Cell Biol. 149:1275-1288(2000).
DR      EMBL; AF116856; AAD24546.2; -.
DR      HSSP; P08709; 1BF9.
DR      InterPro; IPR000152; Asx_hydroxyl.
DR      InterPro; IPR000742; EGF_2.
DR      InterPro; IPR001881; EGF_Ca.
DR      InterPro; IPR001438; EGF_II.
DR      InterPro; IPR006209; EGF_like.
DR      InterPro; IPR003599; Ig.
DR      InterPro; IPR007110; Ig-like.
DR      InterPro; IPR003006; Ig_MHC.
DR      InterPro; IPR001304; Lectin_C.
DR      InterPro; IPR000538; Link.
DR      InterPro; IPR000436; Sushi_SCR_CCP.
DR      Pfam; PF00008; EGF; 2.
DR      Pfam; PF00047; ig; 1.
DR      Pfam; PF00059; lectin_c; 1.
DR      Pfam; PF00084; sushi; 1.
DR      Pfam; PF00193; Xlink; 2.
DR      PRINTS; PR00010; EGFLOOD.
DR      PRINTS; PR01265; LINKMODULE.
DR      ProDom; PD000918; Link; 2.
DR      SMART; SM00032; CCP; 1.
DR      SMART; SM00034; CLECT; 1.
DR      SMART; SM00179; EGF_CA; 1.
DR      SMART; SM00409; IG; 1.
DR      SMART; SM00445; LINK; 2.
DR      PROSITE; PS00010; ASX_HYDROXYL; 1.
DR      PROSITE; PS00615; C_TYPE_LLECTIN_1; 1.
DR      PROSITE; PS50041; C_TYPE_LLECTIN_2; 1.
DR      PROSITE; PS00022; EGF_1; 2.
DR      PROSITE; PS01186; EGF_2; 1.
DR      PROSITE; PS01187; EGF_CA; 1.
DR      PROSITE; PS50835; IG_LIKE; 1.
DR      PROSITE; PS01241; LINK; 2.
KW      EGF-like domain.
SQ      SEQUENCE 1290 AA; 138877 MW; 182BD86D0E40BE78 CRC64;

Query Match      16.7%; Score 185; DB 13; Length 1290;
Best Local Similarity 34.7%; Pred. No. 6.8e-09;
Matches 50; Conservative 18; Mismatches 42; Indels 34; Gaps 8;

Qy      15 QRPCYKVIYFHDTSRRLNFFEEAKEACRRDGGQLVSISEDEQKLIKFIENLLPSDGF 74
|||||
Db      1064 QGHCYR--YF---SRRRSWEDAERDCRRRAGHLTSIHQBEGHGFINSF-----GHENTW 1112
|||||

Qy      75 IGLRRREKQSNSTACQDLYAWTDGSIQFRNWWYVDEPS---CGSEVCVVMY-HQPSAPA 130
|||||
Db      1113 IGLNDRIVEQD-----FQWTDNTGLQYENWRENQPNFAGGEDCVVLVSHE----- 1159
|||||
```

```
Qy      131 GIGGPYMFQWNDRCNMKNPFICK 154
|||||
Db      1160 -IG-----KWNDVPCNYNLPYICK 1177
|||||

RESULT 9
Q61830
ID      Q61830      PRELIMINARY;      PRT; 1456 AA.
AC      Q61830;
DT      01-NOV-1996 (TremBLrel. 01, Created)
DT      01-NOV-1996 (TremBLrel. 01, Last sequence update)
DT      01-MAR-2003 (TremBLrel. 23, Last annotation update)
DE      Macrophage mannose receptor precursor.
GN      MRC1.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      STRAIN=C57BL/6;
RA      Harris N., Rits M., Chang G., Ezekowitz R.B.;
RT      "Characterization of the murine macrophage mannose receptor.";
RL      Blood 80:2363-2373(1992).
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=C57BL/6;
RA      Super M.;
RL      Submitted (APR-1992) to the EMBL/GenBank/DBJ databases.
DR      EMBL; Z11974; CAA78028.1; -.
DR      HSSP; P22897; 1EGG.
DR      MGD; MGI:97142; Mrcl.
DR      InterPro; IPR002353; AntifreezeII.
DR      InterPro; IPR000562; FN_Type_II.
DR      InterPro; IPR001304; Lectin_C.
DR      InterPro; IPR000772; Ricin_B_lectin.
DR      Pfam; PF00040; fn2; 1.
DR      Pfam; PF00059; lectin_c; 8.
DR      Pfam; PF00552; Ricin_B_lectin; 2.
DR      PRINTS; PR00356; ANTIFREEZEII.
DR      PRINTS; PR00013; FNTYPEII.
DR      ProDom; PD000995; FN_Type_II; 1.
DR      SMART; SM00034; CLECT; 8.
DR      SMART; SM00059; FN2; 1.
DR      SMART; SM00458; RICIN; 1.
DR      PROSITE; PS00615; C_TYPE_LLECTIN_1; 6.
DR      PROSITE; PS50041; C_TYPE_LLECTIN_2; 8.
DR      PROSITE; PS00023; FIBRONECTIN_2; 1.
DR      PROSITE; PS50231; RICIN_B_LLECTIN; 1.
KW      Receptor; Signal.
FT      SIGNAL 1 18 POTENTIAL.
FT      CHAIN 19 1456 MACROPHAGE MANNOSE RECEPTOR.
SQ      SEQUENCE 1456 AA; 165065 MW; 4EBD3F1B8619A594 CRC64;

Query Match      16.1%; Score 178.5; DB 11; Length 1456;
Best Local Similarity 25.9%; Pred. No. 3.3e-08;
Matches 53; Conservative 33; Mismatches 66; Indels 53; Gaps 8;

Qy      19 YKVIYFHDTSRRLNFFEEAKEACRRDGGQLVSISEDEQKLIKFIENLLPSDGF 78
|||||
Db      807 YKDYQYFYSKEKETMDNARRCKKNFGDLATIKSESEKFLWKYI-NKNGGQSPYFIGML 865
|||||

Qy      79 RREEKQSNSTACQDLYAWTDGSIQFRNWWYVDEPSCGS--EVCVVMYHQPAPAGIGGPY 136
|||||
Db      866 ISMDKK-----FIWMDGSKVDFVAWATGEPNFANDDENCVTMY-----TNSGF----- 908
|||||

Qy      137 MFQWNDRCNMKNPFICK---YSDEKPAVPSRAEGEBTELTPVLPEETQE----- 185
|||||
Db      909 ---WNDINGYPNNFICQRHNSINATAMP-----TPTTPGCKEGWHLYKNK 954
|||||

Qy      186 -----EDAKTKFKESREAAALNL 202
|||||
```


Db 955 CFKIFGFANEEKKSWQDARQACKGL 979

RESULT 10

Q25199

Q25199 PRELIMINARY; PRT; 1348 AA.

AC Q25199;

DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Tyrosine kinase receptor.

OS Hydra attenuata (Hydra) (Hydra vulgaris).

OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Anthomedusae;

OC Hydridae; Hydra.

OX NCBI_TaxID=6087;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Irvine;

RX MEDLINE=20209407; PubMed=10744720;

RA Reidling J.C., Miller M.A., Steele R.E.;

RT "Sweet Tooth, a Novel Receptor Protein-tyrosine Kinase with C-type

RT Lectin-like Extracellular Domains.";

RL J. Biol. Chem. 275:10323-10330(2000).

DR EMBL; L22612; AAA29218.2; -.

DR HSSP; P11362; 1FGK.

DR InterPro; IPR001304; Lectin C.

DR InterPro; IPR003990; Pancreatias_ac.

DR InterPro; IPR000719; Prot_kinase.

DR InterPro; IPR001245; Tyr_pkinase.

DR Pfam; PF00059; lectin_c; 4.

DR Pfam; PF00069; pkinase; 1.

DR PRINTS; PR01504; PNCREATITSAP.

DR ProDom; PD000001; Prot_kinase; 2.

DR SMART; SM00034; CLECT; 4.

DR SMART; SM00219; TyrcK; 1.

DR PROSITE; PS00615; C_TYPE_LECTIN_1; 2.

DR PROSITE; PS50041; C_TYPE_LECTIN_2; 4.

DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.

DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.

DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.

KW ATP-binding; Kinase; Receptor; Transferase; Tyrosine-protein kinase.

SQ SEQUENCE 1348 AA; 156916 MW; 77D2122093227FEF CRC64;

Query Match 16.0%; Score 177.5; DB 5; Length 1348;

Best Local Similarity 28.0%; Pred. No. 3.8e-08;

Matches 47; Conservative 32; Mismatches 46; Indels 43; Gaps 8;

QY 4 LSGQPVC--RGGTQRPC-----YKVIYFHDTSRLNFEAEAKEACRRDGGQLVSIESTE 55

Db 412 LSHRFICKVKRATNEYCAEGWTSYRIYCYFIYSIEFDWFKSFSSCQNGNLLSIENQEE 471

QY 56 QKLIKFIENLLPSDGD-FWIGLRR-----REEKQSNSTACQDLYAWTDGSIQSQRNWWY 108

Db 472 ----NRFIENDLIKNDKYWIGLKNKIWDYLNKKNR-----FEWSDNTYTOFFNWI 518

QY 109 VDEP--SCGSEVCVVMYHQPSAPAGIGGPPYMFQWDDRCNMKNPFICK 154

Db 519 TNQPDNNGIESCVEMNYN-----GWSDECKEVLNGFICK 553

RESULT 11

Q9Y5P9

Q9Y5P9 PRELIMINARY; PRT; 1479 AA.

AC Q9Y5P9;

DT 01-NOV-1999 (TrEMBLrel. 12, Created)

DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Endocytic receptor Endo180.

GN END0180.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=20148849; PubMed=10683150;

RA Sheikh H., Yarwood H., Ashworth A., Isacke C.;

RT "Endo180, an endocytic recycling glycoprotein related to the

RT macrophage mannose receptor is expressed on fibroblasts, endothelial

RT cells and macrophages and functions as a lectin receptor.";

RL J. Cell Sci. 113:1021-1032(2000).

DR EMBL; AF134838; AAD30280.1; -.

DR HSSP; P02751; 2FN2.

DR InterPro; IPR000562; FN_Type_II.

DR InterPro; IPR001304; Lectin_C.

DR InterPro; IPR000566; Lipocln_cytFABP.

DR InterPro; IPR000772; Ricin_B_lectin.

DR Pfam; PF00040; fn2; 1.

DR Pfam; PF00059; lectin_c; 8.

DR PRINTS; PR00013; FNTYPEII.

DR ProDom; PD000995; FN_Type_II; 1.

DR SMART; SM00034; CLECT; 8.

DR SMART; SM00059; FN2; 1.

DR SMART; SM00458; RICIN; 1.

DR PROSITE; PS00615; C_TYPE_LECTIN_1; 3.

DR PROSITE; PS50041; C_TYPE_LECTIN_2; 8.

DR PROSITE; PS00023; FIBRONECTIN_2; 1.

DR PROSITE; PS00213; LIPOCALIN; 1.

DR PROSITE; PS50231; RICIN_B_LECTIN; 1.

KW Receptor.

SQ SEQUENCE 1479 AA; 166669 MW; 9F4BAF355F036FCE CRC64;

Query Match 16.0%; Score 176.5; DB 4; Length 1479;

Best Local Similarity 30.4%; Pred. No. 5.3e-08;

Matches 51; Conservative 25; Mismatches 57; Indels 35; Gaps 7;

QY 5 SGQPVCRRGGTQRPCYKVIYFHDTSRLNFEAEAKEACRRDGGQLVSIESTEQLIEKFIE 64

Db 385 SWQPF-----QGHCVRL-----QAEXRSWQESKKACLRGGDLVSIHMAELEFITKQIK 434

QY 65 NLLPSDGD-FWIGLRRRREEKQSNSTACQDLYAWTDGSIQSQRNWWYVDEPS---CGSEVCVV 121

Db 435 QEVE---ELWIGL-----NDLKLQMFWSGSLVSFTWHHPFEPNFRDLSLEDCVT 483

QY 122 MYHQPSAPAGIGGPPYMFQWDDRCNMKNPFICKYSDEKPAVPSREAG 169

Db 484 IW---GPEG-----RWNDSPCNQSLPSICKKAGQLSQGAEEHDHG 520

RESULT 12

Q9UBG0

ID Q9UBG0 PRELIMINARY; PRT; 1479 AA.

AC Q9UBG0;

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Urokinase receptor-associated protein UPARAP.

GN KIAA0709.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Behrendt N., Jensen O.N., Engelholm L.H., Mortz E., Mann M., Dano K.;

RT "A urokinase receptor-associated protein with specific collagen-

RT binding properties.";

RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Brain;

RX MEDLINE=98403880; PubMed=9734811;

RA Ishikawa K., Nagase T., Suyama M., Miyajima N., Tanaka A., Kotani H.,

RA Nomura N., Ohara O.;

RT "Prediction of the coding sequences of unidentified human genes. X.

RT The complete sequences of 100 new cDNA clones from brain which can

RT code for large proteins in vitro.";

RL DNA Res. 5:169-176(1998).

DR EMBL; AF107292; AAF14192.1; -.

DR EMBL; AB014609; BAA31684.1; -.

DR HSSP; P02751; 2FN2.

DR InterPro; IPR000562; FN_Type_II.

DR InterPro; IPR001304; Lectin_C.

DR InterPro; IPR000566; Lipocln_cytFABP.

DR InterPro; IPR000772; Ricin_B_lectin.

DR Pfam; PF00040; fn2; 1.

DR Pfam; PF00059; lectin_c; 8.

DR PRINTS; PR00013; FNTYPEII.

DR ProDom; PD000995; FN_Type_II; 1.

DR SMART; SM00034; CLECT; 8.

DR SMART; SM00458; RICIN; 1.

DR PROSITE; PS00615; C_TYPE_LECTIN_1; 3.

DR PROSITE; PS50041; C_TYPE_LECTIN_2; 8.

DR PROSITE; PS00023; FIBRONECTIN_2; 1.

DR PROSITE; PS00213; LIPOCALIN; 1.

DR PROSITE; PS50231; RICIN_B_LECTIN; 1.

KW Kinase; Receptor.

SQ SEQUENCE 1479 AA; 166654 MW; C7583EA78E2792D1 CRC64;

Query Match 16.0%; Score 176.5; DB 4; Length 1479;

Best Local Similarity 30.4%; Pred. No. 5.3e-08;

Matches 51; Conservative 25; Mismatches 57; Indels 35; Gaps 7;

QY 5 SGQPVCRGGTQPCYKVIYFHDTSRRLNFEFEAKEACRRDGGQLVSIESTEDEQKLIKIEFIE 64

Db 385 SWQPF-----QGHCYRL-----QA EKRSWQESKKACLRGGDLVSIHMAELEFITKQIK 434

QY 65 NLLPSDGDGFWIGLRRREEKQSNSTACQDLYAWTDGSI SQFRNWWYVDEPS---CGSEVCVV 121

Db 435 QEVE---ELWIGL-----NDLKLQMNFEWSDGSLVSFTHWHPFPNFRDSDLDCVT 483

QY 122 MYHQPSAPAGIGGPFYMFQWDDRCNMKNFKICKYSDEKPAVPSREAEG 169

Db 484 IW----GPEG-----RWNDSPCNQSLPSICKKAGQLSQGAEEEDHG 520

RESULT 13.

Q64449

ID Q64449 PRELIMINARY; PRT; 1479 AA.

AC Q64449;

DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Lectin lambda.

GN MRC2.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=96355501; PubMed=8702911;

RA Wu K., Yuan J., Lasky L.A.;

RT "Characterization of a novel member of the macrophage mannose receptor

RT type C lectin family.";

RL J. Biol. Chem. 271:21323-21330(1996).

DR EMBL; U56734; AAC52729.1; -.

DR HSSP; P02751; 2FN2.

DR MGD; MGI:107818; Mrc2.

DR InterPro; IPR002353; AntifreezeII.

DR InterPro; IPR001128; Cytochrome_P450.

DR InterPro; IPR000562; FN_Type_II.

DR InterPro; IPR001304; Lectin_C.

DR InterPro; IPR000566; Lipocln_cytFABP.

DR InterPro; IPR000772; Ricin_B_lectin.

DR Pfam; PF00040; fn2; 1.

DR Pfam; PF00059; lectin_c; 8.

DR PRINTS; PR00356; ANTIFREEZEII.

DR PRINTS; PR00013; FNTYPEII.

DR ProDom; PD000995; FN_Type_II; 1.

DR SMART; SM00034; CLECT; 8.

DR SMART; SM00059; FN2; 1.

DR SMART; SM00458; RICIN; 1.

DR PROSITE; PS00086; CYTOCHROME_P450; 1.

DR PROSITE; PS00615; C_TYPE_LECTIN_1; 3.

DR PROSITE; PS50041; C_TYPE_LECTIN_2; 8.

DR PROSITE; PS00023; FIBRONECTIN_2; 1.

DR PROSITE; PS00213; LIPOCALIN; 1.

DR PROSITE; PS50231; RICIN_B_LECTIN; 1.

SQ SEQUENCE 1479 AA; 167112 MW; 62D456E10B9B48C1 CRC64;

Query Match 15.4%; Score 170.5; DB 11; Length 1479;

Best Local Similarity 31.4%; Pred. No. 2e-07;

Matches 48; Conservative 22; Mismatches 48; Indels 35; Gaps 7;

QY 5 SGQPVCRGGTQPCYKVIYFHDTSRRLNFEFEAKEACRRDGGQLVSIESTEDEQKLIKIEFIE 64

Db 384 SWQPF-----QGHCYRL-----QA EKRSWQESKKACLRGGDLVSIHMAELEFITKQIK 433

QY 65 NLLPSDGDGFWIGLRRREEKQSNSTACQDLYAWTDGSI SQFRNWWYVDEPS---CGSEVCVV 121

Db 434 QEVE---ELWIGL-----NDLKLQMNFEWSDGSLVSFTHWHPFPNFRDSDLDCVT 482

QY 122 MYHQPSAPAGIGGPFYMFQWDDRCNMKNFKICK 154

Db 483 IW----GPEG-----RWNDSPCNQSLPSICK 504

RESULT 14

Q8C4F8

ID Q8C4F8 PRELIMINARY; PRT; 217 AA.

AC Q8C4F8;

DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Chondroitin sulfate proteoglycan 3 (Fragment).

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Cerebellum;

RX MEDLINE=22354683; PubMed=12466851;

RA The FANTOM Consortium,

RA the RIKEN Genome Exploration Research Group Phase I & II Team;

RT "Analysis of the mouse transcriptome based on functional annotation of

RT 60,770 full-length cDNAs.";

RL Nature 420:563-573(2002).

DR EMBL; AK082298; BAC38458.1; -.

FT NON TER 1

SQ SEQUENCE 217 AA; 25949 MW; 5F1A6A57B05B76D6 CRC64;

Query Match 15.3%; Score 169; DB 11; Length 217;

Best Local Similarity 31.1%; Pred. No. 2.8e-08;

Matches 42; Conservative 16; Mismatches 47; Indels 30; Gaps 5;

QY 23 YFHDTSRRLNFEFEAKEACRRDGGQLVSIESTEDEQKLIKIEFIENLLPSDGDGFWIGLRRREE 82

Db 3 YF----AHRRAWEDAERDCRRRAGHLTSVHSPEEHKFINSF-----GHENSWIGLNDRTV 53

QY 83 KQSNSTACQDLYAWTDGSI SQFRNWWYVDEPS---CGSEVCVVMYHQPSAPAGIGGPFYMFQ 139

Db 54 ERD-----FQWTDNTGLQYENWREKQPDNFFAGGEDCVVMVAHESG-----R 95

QY 140 WNDRCNMKNFKICK 154

Db 96 WNDVPCNYNLPYVCK 110

RESULT 15

Search completed: December 22, 2003, 16:13:30
Job time : 23.9687 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 22, 2003, 16:05:09 ; Search time 6.56161 Seconds
(without alignments)
1476.391 Million cell updates/sec

Title: US-09-887-855-5
Perfect score: 1115
Sequence: 1 ATGRLLSGQPVCRGGTQRPC.....EEDAKKTFKESREAAALNLAY 206

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	574	51.5	273	1	CHOD_MOUSE
2	561.5	50.4	273	1	CHOD_HUMAN
3	182	16.3	1456	1	MANR_HUMAN
4	177	15.9	1268	1	PGCN_MOUSE
5	176	15.8	1321	1	PGCN_HUMAN
6	174.5	15.7	3381	1	PGCV_BOVIN
7	174	15.6	1257	1	PGCN_RAT
8	174	15.6	2738	1	PGCV_RAT
9	174	15.6	3358	1	PGCV_MOUSE
10	174	15.6	3396	1	PGCV_HUMAN
11	171	15.3	3562	1	PGCV_CHICK
12	165	14.8	643	1	CD93_RAT
13	158.5	14.2	2364	1	PGCA_BOVIN
14	158.5	14.2	2415	1	PGCA_HUMAN
15	155.5	13.9	2333	1	PGCA_CANFA
16	154.5	13.9	612	1	LEM2_MOUSE
17	153.5	13.8	644	1	CD93_MOUSE
18	153.5	13.8	2124	1	PGCA_RAT
19	152	13.6	912	1	PGCB_BOVIN
20	151.5	13.6	652	1	CD93_HUMAN
21	151	13.5	197	1	CLF1_HUMAN
22	149	13.4	321	1	FCE2_HUMAN
23	148.5	13.3	2132	1	PGCA_MOUSE
24	147	13.2	883	1	PGCB_MOUSE
25	146	13.1	2109	1	LECG_CHICK
26	145	13.0	158	1	PGCA_HUMAN
27	145	13.0	883	1	PGCB_RAT
28	144.5	13.0	173	1	LEC2_MEGRO
29	144.5	13.0	372	1	LEM1_RAT
30	143.5	12.9	372	1	LEM1_MOUSE
31	141.5	12.7	331	1	FCE2_MOUSE
32	141.5	12.7	549	1	LEM2_RAT
33	140.5	12.6	162	1	LEC3_MEGRO
					Q9cxm0 mus musculus
					Q9h9p2 homo sapien
					P22897 homo sapien
					P55066 mus musculus
					O14594 homo sapien
					P81282 bos taurus
					P55067 rattus norv
					Q9erb4 rattus norv
					P62059 mus musculus
					P13611 homo sapien
					Q90953 gallus gall
					Q9et61 rattus norv
					P13608 bos taurus
					P16112 homo sapien
					Q28343 canis famil
					Q00690 mus musculus
					O89103 mus musculus
					P07897 rattus norv
					Q28062 bos taurus
					Q9npy3 homo sapien
					O75596 homo sapien
					P06734 homo sapien
					Q61282 mus musculus
					Q61361 mus musculus
					P07898 gallus gall
					Q9y9p1 trimeresuru
					P55068 rattus norv
					P17346 megabalanus
					P30836 rattus norv
					P18337 mus musculus
					P20693 mus musculus
					P98105 rattus norv
					P07439 megabalanus

34	140	12.6	370	1	LEM1_BOVIN	P98131 bos taurus
35	139	12.5	248	1	PSPA_HUMAN	P07714 homo sapien
36	139	12.5	283	1	LECA_SARPE	P05047 sarcophaga
37	138.5	12.4	152	1	IXA_TRIFL	P23806 trimeresuru
38	138.5	12.4	202	1	TETN_MOUSE	P43025 mus musculus
39	137.5	12.3	372	1	LEM1_MACMU	Q95198 macaca mula
40	137.5	12.3	372	1	LEM1_PAPHA	Q28768 papio hamad
41	136	12.2	175	1	LITH_BOVIN	P23132 bos taurus
42	135	12.1	165	1	LIT1_MOUSE	P43137 mus musculus
43	134.5	12.1	132	1	ACAL_ANSAN	P83300 anser anser
44	134.5	12.1	372	1	LEM1_PONPY	Q95235 pongo pygma
45	134.5	12.1	485	1	LEM2_BOVIN	P98107 bos taurus

ALIGNMENTS

RESULT 1
CHOD_MOUSE
ID CHOD_MOUSE STANDARD; PRT; 273 AA.
AC Q9CXM0; Q8VI31;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Chondrolectin precursor (Transmembrane protein MT75).
GN CHODL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RA Weng L., Smits P., Hubner R., Wouters J., Merregaert J.;
RT "Mt75, a low expressed c-type lectin gene involving in
chondrogenesis.";
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryonic head;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF311699; AAL50354.1; -


```
DR Pfam; PF00059; lectin_c; 1.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C_TYPE_LLECTIN_1; FALSE_NEG.
DR PROSITE; PS50041; C_TYPE_LLECTIN_2; 1.
KW Lectin; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 273 CHONDROLECTIN.
FT DOMAIN 22 216 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 217 237 POTENTIAL.
FT DOMAIN 238 273 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 35 179 C-TYPE LECTIN.
FT CARBOHYD 86 86 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 273 AA; 30431 MW; F4890AAPB572A311 CRC64;

Query Match 50.4%; Score 561.5; DB 1; Length 273;
Best Local Similarity 60.1%; Pred. No. 5.7e-43;
Matches 101; Conservative 25; Mismatches 35; Indels 7; Gaps 3;

Qy 4 RLLSGQPVCRGGTQPCYKVIYFHDTSRRLNFEAEKACRRDGGQLVSISEDEQKLIK 63
Db 23 RVVSGQKVCFAFKHPCYKMAYPHELSSRVSFQEARLACESEGGVLLSLENAEQKLIES 82

Qy 64 FIENLLP-----SDGDFWIGLRRREKQSNSTACQDLYAWTDGSIQFRNWWYVDEPSCGS 118
Db 83 MLQNLTKPGTGISDGFWIGLWRNGDGT-SGACPDLYQWSDGNSQYRNWYTDEPSCGS 141

Qy 119 EVCVVMYHOPASAPAGIGGYPYMFQWNNDRCMKNFNICKYSDE-KPAVP 165
Db 142 EKCVMYHQPTANPGLGPGYLYQWNNDRCMKNHYICKYEPEINPTAP 189

RESULT 3
MANR_HUMAN STANDARD; PRT; 1456 AA.
AC P22897;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Macrophage mannose receptor precursor (MMR) (CD206 antigen).
GN MRC1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Placenta;
RX MEDLINE=90324192; PubMed=2373685;
RA Taylor M.E., Conary J.T., Lennartz M.R., Stahl P.D., Drickamer K.;
RT "Primary structure of the mannose receptor contains multiple motifs
RT resembling carbohydrate-recognition domains.";
RL J. Biol. Chem. 265:12156-12162(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93052405; PubMed=1294118;
RA Kim S.J., Ruiz N., Bezouska K., Drickamer K.;
RT "Organization of the gene encoding the human macrophage mannose
RT receptor (MRC1).";
RL Genomics 14:721-727(1992).
RN [3]
RP STUDIES ON THE BINDING OF INDIVIDUAL LECTIN DOMAINS.
RX MEDLINE=92112893; PubMed=1730714;
RA Taylor M.E., Bezouska K., Drickamer K.;
RT "Contribution to ligand binding by multiple carbohydrate-recognition
RT domains in the macrophage mannose receptor.";
RL J. Biol. Chem. 267:1719-1726(1992).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 642-788.
RX MEDLINE=20347275; PubMed=10779515;
RA Feinberg H., Park-Snyder S., Kolatkar A.R., Heise C.T., Taylor M.E.,
RA Weis W.I.;
RT "Structure of a C-type carbohydrate recognition domain from the
RT macrophage mannose receptor.";
```

```
RL J. Biol. Chem. 275:21539-21548(2000).
CC -!- FUNCTION: MEDIATES THE ENDOCYTOSIS OF GLYCOPROTEINS BY
CC MACROPHAGES, IN SEVERAL RECOGNITION AND UPTAKE PROCESSES.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- MISCELLANEOUS: CRDS 1-3 HAVE AT MOST VERY WEAK AFFINITY FOR
CC CARBOHYDRATE. CRD 4 SHOWS THE HIGHEST AFFINITY BINDING AND HAS
CC MULTISPECIFICITY FOR A VARIETY OF MONOSACCHARIDES. AT LEAST 3 CRDS
CC (4, 5, AND 7) ARE REQUIRED FOR HIGH AFFINITY BINDING AND
CC ENDOCYTOSIS OF MULTIVALENT GLYCOCONJUGATES.
CC -!- SIMILARITY: Contains 8 C-type lectin family domains.
CC -!- SIMILARITY: Contains 1 ricin B-type lectin domain.
CC -!- DATABASE: NAME=PROW; NOTE=PROW 2:85-89(2001);
CC WWW="http://www.ncbi.nlm.nih.gov/prow/guide/1644341535_g.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; J05550; AAA59868.1; -
CC EMBL; M93221; AAA60389.1; -
CC EMBL; M93192; AAA60389.1; JOINED.
CC EMBL; M93193; AAA60389.1; JOINED.
CC EMBL; M93194; AAA60389.1; JOINED.
CC EMBL; M93195; AAA60389.1; JOINED.
CC EMBL; M93196; AAA60389.1; JOINED.
CC EMBL; M93197; AAA60389.1; JOINED.
CC EMBL; M93198; AAA60389.1; JOINED.
CC EMBL; M93199; AAA60389.1; JOINED.
CC EMBL; M93200; AAA60389.1; JOINED.
CC EMBL; M93201; AAA60389.1; JOINED.
CC EMBL; M93202; AAA60389.1; JOINED.
CC EMBL; M93203; AAA60389.1; JOINED.
CC EMBL; M93204; AAA60389.1; JOINED.
CC EMBL; M93205; AAA60389.1; JOINED.
CC EMBL; M93206; AAA60389.1; JOINED.
CC EMBL; M93207; AAA60389.1; JOINED.
CC EMBL; M93208; AAA60389.1; JOINED.
CC EMBL; M93209; AAA60389.1; JOINED.
CC EMBL; M93210; AAA60389.1; JOINED.
CC EMBL; M93211; AAA60389.1; JOINED.
CC EMBL; M93212; AAA60389.1; JOINED.
CC EMBL; M93213; AAA60389.1; JOINED.
CC EMBL; M93214; AAA60389.1; JOINED.
CC EMBL; M93215; AAA60389.1; JOINED.
CC EMBL; M93216; AAA60389.1; JOINED.
CC EMBL; M93217; AAA60389.1; JOINED.
CC EMBL; M93218; AAA60389.1; JOINED.
CC EMBL; M93219; AAA60389.1; JOINED.
CC EMBL; M93220; AAA60389.1; JOINED.
CC PIR; A36563; A36563.
CC PDB; 1EGG; 30-AUG-00.
CC PDB; 1EGI; 30-AUG-00.
CC Genew; HGNC:7228; MRC1.
CC MIM; 153618; -
CC GO; GO:0005887; C:integral to plasma membrane; TAS.
CC GO; GO:0005337; F:mannose binding activity; TAS.
CC GO; GO:0004872; F:receptor activity; TAS.
CC GO; GO:0006898; P:receptor mediated endocytosis; TAS.
CC InterPro; IPR002353; AntifreezeII.
CC InterPro; IPR000562; FN_Type_II.
CC InterPro; IPR001304; Lectin_C.
CC InterPro; IPR000772; Ricin_B_lectin.
CC Pfam; PF00040; fn2; 1.
CC Pfam; PF00059; lectin_c; 8.
CC Pfam; PF00652; Ricin_B_lectin; 2.
CC PRINTS; PR00013; FNTYPEII.
CC PRINTS; PR00356; ANTIFREEZEII.
CC ProDom; PD000995; FN_Type_II; 1.
CC SMART; SM00034; CLECT; 8.
```

DR SMART; SM00059; FN2; 1.
DR SMART; SM00458; RICIN; 1.
DR PROSITE; PS00615; C_TYPE_LECTIN_1; 6.
DR PROSITE; PS00041; C_TYPE_LECTIN_2; 8.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00231; RICIN_B_LECTIN; 1.
KW Receptor; Signal; Calcium-binding; Transmembrane; Repeat;
KW Glycoprotein; Endocytosis; Macrophage; Lectin; Antigen; 3D-structure.
FT SIGNAL 1 18
FT CHAIN 19 1456
FT DOMAIN 19 1383
FT TRANSMEM 1384 1411
FT DOMAIN 1412 1456
FT DOMAIN 22 142
FT DOMAIN 157 212
FT DOMAIN 216 344
FT DOMAIN 360 490
FT DOMAIN 502 629
FT DOMAIN 644 781
FT DOMAIN 805 926
FT DOMAIN 943 1083
FT DOMAIN 1100 1216
FT DOMAIN 1228 1359
FT DISULFID 646 659
FT DISULFID 680 777
FT DISULFID 753 769
FT CARBOHYD 104 104
FT CARBOHYD 344 344
FT CARBOHYD 529 529
FT CARBOHYD 926 926
FT CARBOHYD 930 930
FT CARBOHYD 1160 1160
FT CARBOHYD 1205 1205
FT CARBOHYD 1311 1311
FT TURN 648 649
FT STRAND 651 652
FT TURN 654 655
FT STRAND 658 663
FT HELIX 667 669
FT STRAND 671 671
FT HELIX 673 683
FT TURN 684 684
FT STRAND 686 687
FT HELIX 693 705
FT TURN 706 707
FT TURN 709 710
FT STRAND 712 718
FT TURN 723 724
FT STRAND 727 727
FT TURN 729 730
FT STRAND 733 733
FT TURN 741 742
FT HELIX 746 748
FT STRAND 752 757
FT TURN 758 761
FT STRAND 764 768
FT TURN 769 770
FT STRAND 773 780
FT TURN 781 782
SQ SEQUENCE 1456 AA; 166011 MW; 264E5AF3C576A5E3 CRC64;

Query Match 16.3%; Score 182; DB 1; Length 1456;
Best Local Similarity 25.5%; Pred. No. 3.4e-08;
Matches 52; Conservative 37; Mismatches 61; Indels 54; Gaps 8;

QY 21 YKVIYFHDTSRRLNFEAEKACRRDGGQLVSIQSEDEQKLEKFIENLLPSDGFWIGLR 80
DB 807 YKDYQYYSKEKETMDNARAFCKRNFGDLVSIQSEKFLWKYV-NRNDQAQSAFYIGLL 865
QY 81 RREEKQSNSTACQDLVYAWTDGSIQFRNWWYVDEPSCS--EVCVVMYHQPSAPAGIGGPY 138
DB 866 ISLDKK-----FAWMDGSKVDYVSWATGEPNANEDENCVTMY-----SNSGF---- 908

QY 139 MFQWDDRCNMKNFKICKYSDEK----PAVPSREAGEETELTTPVLPEETQE----- 187
DB 909 ---WNDINCGYPNAFICQHNSSINATTVM-----TMPSPVSGCKEGWNFYSN 954
QY 188 -----EDAKKTPKESREAAAL 202
DB 955 KCFKIFGFMEERKNWQEARAKACI 978

RESULT 4
PGCN MOUSE
ID_PGCN_MOUSE STANDARD; PRT; 1268 AA.
AC P55066;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Neurocan core protein precursor (Chondroitin sulfate proteoglycan 3).
GN CSPG3 OR NCAN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C; TISSUE=Brain;
RX MEDLINE=96039250; PubMed=7490074;
RA Rauch U., Grimpe B., Kulbe G., Arnold-Ammer I., Beier D.,
RA Faessler R.;
RT "Structure and chromosomal localization of the mouse neurocan gene.";
RL Genomics 28:405-410(1995).
CC -!- FUNCTION: May modulate neuronal adhesion and neurite growth during
CC development by binding to neural cell adhesion molecules (NG-CAM
CC and N-CAM). Chondroitin sulfate proteoglycan; binds to hyaluronic
CC acid.
CC -!- TISSUE SPECIFICITY: BRAIN.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -!- SIMILARITY: Contains 2 EGF-like domains.
CC -!- SIMILARITY: Contains 2 link domains.
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC -!- SIMILARITY: Contains 1 Sushi (SCR) domain.
CC -!- SIMILARITY: BELONGS TO THE AGGRECAN/VERSICAN PROTEOGLYCAN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; X84727; CAA59216.1; -.
DR PIR; S52781; S52781.
DR HSSP; P00740; 1EDM.
DR MGD; MGI:104694; Cspg3.
DR InterPro; IPR002353; Antifreeze1.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR000538; Link.
DR InterPro; IPR000436; Sushi_SCR_CCP.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF00059; lectin_c; 1.
DR Pfam; PF00084; sushi; 1.
DR Pfam; PF00193; xlink; 2.
DR PRINTS; PR01265; LINKMODULE.
DR PRINTS; PR00356; ANTIFREEZE1.
DR ProDom; PD000918; Link; 2.


```
DR PROSITE; PS01187; EGF_CA; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR PROSITE; PS01241; LINK; 2.
KW Glycoprotein; Hyaluronic acid; Proteoglycan; Immunoglobulin domain;
KW EGF-like domain; Calcium; Repeat; Lectin; Sushi; Signal.
FT SIGNAL 1 22
FT CHAIN 23 1321 NEUROCAN CORE PROTEIN.
FT DOMAIN 38 153 IG-LIKE V-TYPE.
FT DOMAIN 159 254 LINK 1.
FT DOMAIN 260 356 LINK 2.
FT DOMAIN 1008 1044 EGF-LIKE 1.
FT DOMAIN 1046 1082 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 1084 1213 C-TYPE LECTIN.
FT DOMAIN 1214 1272 SUSHI.
FT DISULFID 59 140 BY SIMILARITY.
FT DISULFID 182 253 BY SIMILARITY.
FT DISULFID 206 227 BY SIMILARITY.
FT DISULFID 280 355 BY SIMILARITY.
FT DISULFID 304 325 BY SIMILARITY.
FT DISULFID 1012 1023 BY SIMILARITY.
FT DISULFID 1017 1032 BY SIMILARITY.
FT DISULFID 1034 1043 BY SIMILARITY.
FT DISULFID 1088 1099 BY SIMILARITY.
FT DISULFID 1116 1208 BY SIMILARITY.
FT DISULFID 1184 1200 BY SIMILARITY.
FT DISULFID 1215 1258 BY SIMILARITY.
FT DISULFID 1244 1271 BY SIMILARITY.
FT CARBOHYD 122 122 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 340 340 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1026 1026 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1223 1223 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 1234 1234 N -> Y (IN REF. 2).
FT CONFLICT 1254 1254 V -> A (IN REF. 2).
FT CONFLICT 1282 1282 G -> R (IN REF. 2).
SQ SEQUENCE 1321 AA; 142972 MW; 2EF47F823DB980B8 CRC64;

Query Match 15.8%; Score 176; DB 1; Length 1321;
Best Local Similarity 31.5%; Pred. No. 1e-07;
Matches 45; Conservative 18; Mismatches 48; Indels 32; Gaps 6;

QY 17 QRPCYKVIYFHTSRRUNFEEAKEACRRDGGQLVSIIESEDEQKLIKFIENLPSDGDWF 76
Db 1096 QGHCYR--YF---AHRAWEDAECDCRRRSGHLTSVHSPEHSFINSF-----GHENTW 1144

QY 77 IGLRRREEKQSNSTACQLYAWTDGSIQFRNWWYVDEPS---CGSEVCVVMYHQSAPAG 133
Db 1145 IGLNDRIVERD-----FQWTDNTGLQFENWRENQPDNFFAGGEDCVVMVAHESG--- 1193

QY 134 IGGPYMFQWDDRCNMKNFNICK 156
Db 1194 -----RWNDVPCYNLPLVCK 1209

RESULT 6
PGCV_BOVIN
ID -PGCV_BOVIN STANDARD; PRT; 3381 AA.
AC P81282; O77609; O77610; O77611; O77612;
DT 15-DEC-1998 (Rel. 37, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Versican core protein precursor (Large fibroblast proteoglycan)
DE (Chondroitin sulfate proteoglycan core protein 2) (PG-M) (Glial
DE hyaluronate-binding protein) (GHAP).
GN CPSC2.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS VO; V1; V2 AND V3).
RC TISSUE=Forebrain;
RX MEDLINE=98288320; PubMed=9624174;
```

```
RA Schmalfeldt M., Dours-Zimmermann M.T., Winterhalter K.H.,
RA Zimmermann D.R.;
RT "Versican V2 is a major extracellular matrix component of the mature
RT bovine brain.";
RL J. Biol. Chem. 273:15758-15764 (1998).
RN [2]
RP SEQUENCE OF 21-53; 78-96; 226-250; 262-277; 295-306; 314-324; 329-331
RP AND 342-348.
RC TISSUE=Spinal cord;
RX MEDLINE=92062692; PubMed=1720020;
RA Perides G., Biviano F., Bignami A.;
RT "Interaction of a brain extracellular matrix protein with hyaluronic
RT acid.";
RL Biochim. Biophys. Acta 1075:248-258 (1991).
CC -!- FUNCTION: May play a role in intercellular signaling and in
CC connecting cells with the extracellular matrix. May take part in
CC the regulation of cell motility, growth and differentiation. Binds
CC hyaluronic acid.
CC -!- SUBUNIT: Interacts with FBLN1 (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=4;
CC Comment=Additional isoforms seem to exist;
CC Name=V0;
CC IsoId=P81282-1; Sequence=Displayed;
CC Name=V1;
CC IsoId=P81282-2; Sequence=VSP_003078, VSP_003079;
CC Name=V2;
CC IsoId=P81282-3; Sequence=VSP_003080;
CC Name=V3;
CC IsoId=P81282-4; Sequence=VSP_003078, VSP_003081;
CC -!- TISSUE SPECIFICITY: Cerebral white matter. V0 and V1 are expressed
CC in the central nervous system, and in a number of mesenchymal and
CC epithelial tissues; the major isoform V2 is restricted to the
CC central nervous system.
CC -!- DEVELOPMENTAL STAGE: Disappears after the cartilage development
CC (By similarity).
CC -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -!- SIMILARITY: Contains 2 link domains.
CC -!- SIMILARITY: Contains 2 EGF-like domains.
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC -!- SIMILARITY: Contains 1 Sushi (SCR) domain.
CC -!- SIMILARITY: BELONGS TO THE AGGRECAN/VERSICAN PROTEOGLYCAN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF060456; AAC24358.1; -
DR EMBL; AF060457; AAC24359.1; -
DR EMBL; AF060458; AAC24360.1; -
DR EMBL; AF060459; AAC24361.1; -
DR PIR; T14274; T14274.
DR PIR; T42389; T42389.
DR HSSP; P01132; 1EPG.
DR InterPro; IPR000152; Aax_hydroxyl.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR000538; Link.
DR InterPro; IPR000436; Sushi_SCR_CCP.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF00059; lectin_c; 1.
DR Pfam; PF00084; sushi; 1.
```

DR Pfam; PF00193; Xlink; 2.
DR PRINTS; PRO1265; LINKMODULE.
DR PRINTS; PRO0356; ANTIFREEZEII.
DR ProDom; PD000918; Link; 2.
DR SMART; SM00032; CCP; 1.
DR SMART; SM00034; CLECT; 1.
DR SMART; SM00179; EGF_CA; 1.
DR SMART; SM00409; IG_1.
DR SMART; SM00445; LINK; 2.
DR PROSITE; PS00010; ASX_HYDROXYL; 1.
DR PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR PROSITE; PS00041; C_TYPE_LECTIN_2; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01187; EGF_CA; 1.
DR PROSITE; PS00835; IG_LIKE; 1.
DR PROSITE; PS01241; LINK; 2.
KW Glycoprotein; Proteoglycan; Lectin; Extracellular matrix; Sushi;
KW Signal; Repeat; EGF-like domain; Calcium; Immunoglobulin domain;
KW Hyaluronic acid; Alternative splicing.
FT SIGNAL 1 20
FT CHAIN 21 3381
FT DOMAIN 21 147
FT DOMAIN 168 245
FT DOMAIN 266 347
FT DOMAIN 349 1336
FT DOMAIN 1337 3074
FT DOMAIN 3074 3110
FT DOMAIN 3112 3148
FT DOMAIN 3161 3275
FT DOMAIN 3280 3338
FT DISULFID 44 131
FT DISULFID 173 244
FT DISULFID 197 218
FT DISULFID 271 346
FT DISULFID 295 316
FT DISULFID 3078 3089
FT DISULFID 3083 3098
FT DISULFID 3100 3109
FT DISULFID 3116 3127
FT DISULFID 3121 3136
FT DISULFID 3138 3147
FT DISULFID 3154 3165
FT DISULFID 3182 3274
FT DISULFID 3250 3266
FT DISULFID 3281 3324
FT DISULFID 3310 3337
FT CARBOHYD 57 57
FT CARBOHYD 331 331
FT CARBOHYD 352 352
FT CARBOHYD 817 817
FT CARBOHYD 965 965
FT CARBOHYD 1017 1017
FT CARBOHYD 1333 1333
FT CARBOHYD 1393 1393
FT CARBOHYD 1437 1437
FT CARBOHYD 1463 1463
FT CARBOHYD 1653 1653
FT CARBOHYD 1974 1974
FT CARBOHYD 2045 2045
FT CARBOHYD 2074 2074
FT CARBOHYD 2103 2103
FT CARBOHYD 2263 2263
FT CARBOHYD 2290 2290
FT CARBOHYD 2356 2356
FT CARBOHYD 2623 2623
FT CARBOHYD 2641 2641
FT CARBOHYD 2919 2919
FT CARBOHYD 3052 3052
FT CARBOHYD 3354 3354
FT CARBOHYD 3364 3364
FT VARSPLIC 349 349

FT VARSPLIC 350 1336 /FTid=VSP 003078.
FT Missing (in isoform V1).
FT /FTid=VSP 003079.
FT Missing (in isoform V2).
FT /FTid=VSP 003080.
FT Missing (in isoform V3).
FT /FTid=VSP 003081.
FT MISSING (IN REF. 2).
FT CONFLICT 25 25 MISSING (IN REF. 2).
FT CONFLICT 51 51 MISSING (IN REF. 2).
FT CONFLICT 89 89 N -> D (IN REF. 2).
FT CONFLICT 96 96 Q -> D (IN REF. 2).
FT CONFLICT 346 346 C -> R (IN REF. 2).
SQ SEQUENCE 3381 AA; 369984 MW; F09716FA778D459 CRC64;

Query Match 15.7%; Score 174.5; DB 1; Length 3381;
Est Local Similarity 25.5%; Pred. No. 4.5e-07;
Matches 59; Conservative 26; Mismatches 71; Indels 75; Gaps 11;

QY 17 QRPCYKVIYFHDTSRRRLNFEFAKEACRRDGGQLVLSIESDEQKLIKFIENLLPSDGD- 75
Db 3162 QGQCYK--YF--AHRRTWDAARECRLOGAHLTSILSHEEQMFVNRV-----GHYQ 3209
QY 76 WIGLRRREKQSNSTACQDLYAWTDGSIQFRNWWYDEP-----SCGSEVCVVMYHQSAP 131
Db 3210 WIGL-----NDKMFEDHFRWTDGSLQYENWRPNQDPSFFSTGEDCVIWHENG-- 3259
QY 132 AGIGGPMFQWDDRCNMKNFICKYS-----DEKPAVPSREAEGE----- 172
Db 3260 -----QWNVPCNYHLTYCKKGTGTVACGQPPVVENAKTFGKMKPRYEINSLIRYHC 3310
QY 173 -----ETELTT-----PVL-----PEETQEEDAKTKFKESREAAALN 203
Db 3311 KDGFIQRLPTIRCLNGRWAMPKITCLNPSAYQRTYSKKYFKNSSSAKDN 3361

RESULT 7
PGCN_RAT
ID_PGCN_RAT STANDARD; PRT; 1257 AA.
AC P55067;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Neurocan core protein precursor (Chondroitin sulfate proteoglycan 3)
DE (245 kDa early postnatal core glycoprotein) [Contains: 150 kDa adult
core glycoprotein].
GN CSPG3 OR NCAN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=Sprague-Dawley; TISSUE=Brain;
RX MEDLINE=92406907; PubMed=1326557;
RA Rauch U., Karthikeyan L., Maurel P., Margolis R.U., Margolis R.K.;
RT "Cloning and primary structure of neurocan, a developmentally
regulated, aggregating chondroitin sulfate proteoglycan of brain.";
RL J. Biol. Chem. 267:19536-19547(1992).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=94230574; PubMed=7513709;
RA Friedlander D.R., Milev P., Karthikeyan L., Margolis R.K.,
RA Margolis R.U., Grumet M.;
RT "The neuronal chondroitin sulfate proteoglycan neurocan binds to the
neural cell adhesion molecules Ng-CAM/L1/NILE and N-CAM, and inhibits
neural adhesion and neurite outgrowth.";
RL J. Cell Biol. 125:669-680(1994).
CC -!- FUNCTION: May modulate neuronal adhesion and neurite growth during
development by binding to neural cell adhesion molecules (NG-CAM
and N-CAM). Chondroitin sulfate proteoglycan; binds to hyaluronic
acid.
CC -!- TISSUE SPECIFICITY: EARLY POSTNATAL AND ADULT BRAIN; NOT EXPRESSED
IN KIDNEY, LUNG, LIVER AND MUSCLE.

[4]
SEQUENCE OF 2535-2738 FROM N.A.
STRAIN=Sprague-Dawley; TISSUE=Lung;
Blomberg L.A., Chan W.-Y., Clerch L., Massaro D.;
"Molecular cloning and characterization of two developmentally
regulated genes in rat lung.";
Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
-!- FUNCTION: May play a role in intercellular signaling and in
connecting cells with the extracellular matrix. May take part in
the regulation of cell motility, growth and differentiation. Binds
hyaluronic acid.
-!- SUBUNIT: Interacts with FBLN1 (By similarity).
-!- SUBCELLULAR LOCATION: Secreted; extracellular matrix.
-!- ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=3;
Comment=Additional isoforms seem to exist;
Name=V0;
IsoId=Q9ERB4-1; Sequence=Displayed;
Name=V3;
IsoId=Q9ERB4-2; Sequence=VSP_003091;
Name=Vint;
IsoId=Q9ERB4-3; Sequence=VSP_003092;
-!- TISSUE SPECIFICITY: In kidney is expressed in the papillary area,
but not in glomeruli.
-!- DEVELOPMENTAL STAGE: Disappears after the cartilage development
(By similarity).
-!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
-!- SIMILARITY: Contains 2 link domains.
-!- SIMILARITY: Contains 2 EGF-like domains.
-!- SIMILARITY: Contains 1 C-type lectin family domain.
-!- SIMILARITY: Contains 1 Sushi (SCR) domain.
-!- SIMILARITY: BELONGS TO THE AGGRECAN/VERSICAN PROTEOGLYCAN FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; AF062402; AAC40166.1; -.
EMBL; U75306; AAB51125.1; -.
EMBL; AF084544; AAC48544.1; -.
EMBL; AF072892; AAC26116.1; -.
EMBL; AY007691; AAG16631.1; -.
HSSP; P01132; 1EPG.
InterPro; IPR000152; Asx_hydroxyl.
InterPro; IPR000742; EGF_2.
InterPro; IPR001881; EGF_Ca.
InterPro; IPR006209; EGF_like.
InterPro; IPR007110; Ig-like.
InterPro; IPR003599; Ig.
InterPro; IPR001304; Lectin_C.
InterPro; IPR000538; Link.
InterPro; IPR000436; Sushi_SCR_CCP.
PRINTS; PR01265; LINKMODULE.
PRINTS; PR00356; ANTIFREEZEII.
ProDom; PD000918; Link; 2.
SMART; SM00032; CCP; 1.
SMART; SM00034; CLECT; 1.
SMART; SM00179; EGF_CA; 1.
SMART; SM00409; IG; 1.
SMART; SM00445; LINK; 2.
PROSITE; PS00010; ASX_HYDROXYL; 1.
PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
PROSITE; PS00022; EGF_1; 2.
PROSITE; PS01186; EGF_2; 1.
PROSITE; PS01187; EGF_CA; 1.
PROSITE; PS50835; IG_LINK; 1.
PROSITE; PS01241; LINK; 2.
Glycoprotein; Proteoglycan; Lectin; Extracellular matrix; Sushi;

KW	Signal; Repeat; EGF-like domain; Calcium; Immunoglobulin domain;
KW	Hyaluronic acid; Alternative splicing.
FT	SIGNAL 1 20 POTENTIAL.
FT	CHAIN 21 2738 VERSICAN CORE PROTEIN.
FT	NON CONS 348 349
FT	DOMAIN 21 146 IG-LIKE V-TYPE.
FT	DOMAIN 167 244 LINK 1.
FT	DOMAIN 265 346 LINK 2.
FT	DOMAIN <349 695 GAG-ALPHA
FT	(GLUCOSAMINOGLYCAN ATTACHMENT DOMAIN).
FT	DOMAIN 696 2431 GAG-BETA.
FT	DOMAIN 2431 2467 EGF-LIKE 1.
FT	DOMAIN 2469 2505 EGF-LIKE 2, CALCIUM-BINDING.
FT	DOMAIN 2518 2632 C-TYPE LECTIN.
FT	DOMAIN 2637 2695 SUSHI.
FT	DISULFID 44 130 BY SIMILARITY.
FT	DISULFID 172 243 BY SIMILARITY.
FT	DISULFID 196 217 BY SIMILARITY.
FT	DISULFID 270 345 BY SIMILARITY.
FT	DISULFID 294 315 BY SIMILARITY.
FT	DISULFID 2435 2446 BY SIMILARITY.
FT	DISULFID 2440 2455 BY SIMILARITY.
FT	DISULFID 2457 2466 BY SIMILARITY.
FT	DISULFID 2473 2484 BY SIMILARITY.
FT	DISULFID 2478 2493 BY SIMILARITY.
FT	DISULFID 2495 2504 BY SIMILARITY.
FT	DISULFID 2511 2522 BY SIMILARITY.
FT	DISULFID 2539 2631 BY SIMILARITY.
FT	DISULFID 2607 2623 BY SIMILARITY.
FT	DISULFID 2638 2681 BY SIMILARITY.
FT	DISULFID 2667 2694 BY SIMILARITY.
FT	CARBOHYD 57 57 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 330 330 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 692 692 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 758 758 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 802 802 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 805 805 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 1257 1257 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 1435 1435 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 1633 1633 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 1660 1660 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 1684 1684 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 1738 1738 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 1848 1848 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 2004 2004 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 2409 2409 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 2711 2711 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 2721 2721 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	VARSPLIC 349 Missing (in isoform V3).
FT	/FTid=VSP_003091.
FT	VARSPPLIC 2697 2738 PSAYQRTYSKKYLKSSSVKDINSINTSKHEHRWSRRWQETR
FT	R -> RKWFSRKNQPCCFNKY (in isoform Vint).
FT	/FTid=VSP_003092.
FT	CONFLICT 2535 2539 AEREC -> NSARG (IN REF. 4).
FT	SEQUENCE 2738 AA; 300004 MW; 12CA626D58BD8C6A CRC64;
SQ	
Query Match 15.6%; Score 174; DB 1; Length 2738;	
Best Local Similarity 28.5%; Pred. No. 3.9e-07;	
Matches 47; Conservative 23; Mismatches 55; Indels 40; Gaps 8;	
QY	17 QRPCYKVIFYFDTSRRLNFEAEAKEARRDGGQLVSIESEDEQKLIEFIENLLPSDGDF- 75
Db	2519 QGQCYN--YF---AHRRTWDAARECRLOGAHLTSLSHEEQMFVN RV-----GHDIYQ 2566
QY	76 WIGLRRREEKQSNTACQDLIAWTGDSISQFRNWWYVDEP----SCGSEVCVMYHQPSAP 131
Db	2567 WIGL-----NDKMFEHDFRWTDGSALQYENWRPNQDPDSFFSAGEDCVVIWHENG-- 2616
QY	132 AGIGGPYMFQNDDRCNMKNNFICKYS----DEKPAVPSPREAGEE 172
Db	2617 -----QWNDRVPCNYHLTYTCCKGTGVACGPPPVVENAKTFGK 2652


```
RESULT 9
PGCV MOUSE
ID PGCV_MOUSE STANDARD; PRT: 3358 AA.
AC Q62059; Q62058; Q9CUU0;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Versican core protein precursor (large fibroblast proteoglycan)
DE (Chondroitin sulfate proteoglycan core protein 2) (PG-M).
DE CSPG2.
GN Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS V0; V1 AND V2).
RC STRAIN=C57BL/6, and Swiss Webster; TISSUE=Brain;
RX MEDLINE=95122551; PubMed=7822336;
RA Ito K., Shinomura T., Zako M., Ujita M., Kimata K.;
RT "Multiple forms of mouse PG-M, a large chondroitin sulfate
RT proteoglycan generated by alternative splicing.";
RL J. Biol. Chem. 270:958-965(1995).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM V3).
RC STRAIN=C57BL/6;
RX MEDLINE=95181355; PubMed=7876137;
RA Zako M., Shinomura T., Ujita M., Ito K., Kimata K.;
RT "Expression of PG-M(V3), an alternatively spliced form of PG-M
RT without a chondroitin sulfate attachment in region in mouse and human
RT tissues.";
RL J. Biol. Chem. 270:3914-3918(1995).
RN [3]
RP SEQUENCE OF 1-1692 FROM N.A. (ISOFORM V1).
RC STRAIN=C57BL/6J; TISSUE=Skin;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombarts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [4]
RP INTERACTION WITH FBLN1.
RX PubMed=10400671;
RA Asperger A., Adam S., Kostka G., Timpl R., Heinegaard D.;
RT "Fibulin-1 is a ligand for the C-type lectin domains of aggrecan and
RT versican.";
RL J. Biol. Chem. 274:20444-20449(1999).
CC -!- FUNCTION: May play a role in intercellular signaling and in
CC connecting cells with the extracellular matrix. May take part in
CC the regulation of cell motility, growth and differentiation. Binds
CC hyaluronic acid.
CC -!- SUBUNIT: Interacts with FBLN1.
CC -!- CELLULAR LOCATION: Secreted; extracellular matrix.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=4;
CC Comment=Additional isoforms seem to exist;
CC Name=V0;
CC IsoId=Q62059-1; Sequence=Displayed;
```

```
CC Name=V1;
CC IsoId=Q62059-2; Sequence=VSP_003087, VSP_003088;
CC Name=V2;
CC IsoId=Q62059-3; Sequence=VSP_003089;
CC Name=V3;
CC IsoId=Q62059-4; Sequence=VSP_003087, VSP_003090;
CC TISSUE SPECIFICITY: V2 is found only in brain.
CC -!- DEVELOPMENTAL STAGE: Disappears after the cartilage development.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -!- SIMILARITY: Contains 2 link domains.
CC -!- SIMILARITY: Contains 2 EGF-like domains.
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC -!- SIMILARITY: Contains 1 Sushi (SCR) domain.
CC -!- SIMILARITY: BELONGS TO THE AGGREGAN/VERSICAN PROTEOGLYCAN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D16263; BAA03796.1; -
CC EMBL; D28599; -; NOT ANNOTATED_CDS.
CC EMBL; D32040; BAA06802.1; -
CC EMBL; AK014525; BAB29411.1; -
CC HSSP; P01132; 1EPG.
CC MGD; MGI:102889; Cspg2.
CC InterPro; IPR000152; Asx_hydroxyl.
CC InterPro; IPR000742; EGF_2.
CC InterPro; IPR001881; EGF_Ca.
CC InterPro; IPR006209; EGF_like.
CC InterPro; IPR007110; Ig_Like.
CC InterPro; IPR003599; Ig.
CC InterPro; IPR003006; Ig_MHC.
CC InterPro; IPR001304; Lectin_C.
CC InterPro; IPR000538; Link.
CC InterPro; IPR000436; Sushi_SCR_CCP.
CC Pfam; PF00008; EGF; 2.
CC Pfam; PF00047; ig; 1.
CC Pfam; PF00059; lectin_c; 1.
CC Pfam; PF00084; sushi; 1.
CC Pfam; PF00193; Xlink; 2.
CC PRINTS; PR01265; LINKMODULE.
CC PRODOM; PD000918; Link; 2.
CC SMART; SM00032; CCP; 1.
CC SMART; SM00034; CLECT; 1.
CC SMART; SM00179; EGF_CA; 1.
CC SMART; SM00409; IG; 1.
CC SMART; SM00445; LINK; 2.
CC PROSITE; PS00010; ASX_HYDROXYL; 1.
CC PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
CC PROSITE; PS00041; C_TYPE_LECTIN_2; 1.
CC PROSITE; PS00022; EGF_1; 2.
CC PROSITE; PS01186; EGF_2; 1.
CC PROSITE; PS01187; EGF_CA; 1.
CC PROSITE; PS00835; IG_LIKE; 1.
CC PROSITE; PS01241; LINK; 2.
CC Glycoprotein; Proteoglycan; Lectin; Extracellular matrix; Sushi;
CC Signal; Repeat; EGF-like domain; Calcium; immunoglobulin domain;
CC Hyaluronic acid; Alternative splicing.
CC SIGNAL 1 20
CC CHAIN 21 3358
CC DOMAIN 21 146
CC DOMAIN 167 244
CC DOMAIN 265 346
CC DOMAIN 348 1308
CC (GLUCOSAMINOGLYCAN ATTACHMENT DOMAIN).
CC GAG-BETA.
CC 1309 3052
CC 3052 3088
CC 3090 3126
CC 3139 3253
CC EGF-LIKE 1.
CC EGF-LIKE 2.
CC CALCIUM-BINDING (POTENTIAL).
CC C-TYPE LECTIN.
```

FT	DOMAIN	3258	3316	SUSH1.
FT	DISULFID	44	130	BY SIMILARITY.
FT	DISULFID	172	243	BY SIMILARITY.
FT	DISULFID	196	217	BY SIMILARITY.
FT	DISULFID	270	333	BY SIMILARITY.
FT	DISULFID	294	315	BY SIMILARITY.
FT	DISULFID	3056	3067	BY SIMILARITY.
FT	DISULFID	3061	3076	BY SIMILARITY.
FT	DISULFID	3078	3087	BY SIMILARITY.
FT	DISULFID	3094	3105	BY SIMILARITY.
FT	DISULFID	3099	3114	BY SIMILARITY.
FT	DISULFID	3116	3125	BY SIMILARITY.
FT	DISULFID	3132	3143	BY SIMILARITY.
FT	DISULFID	3160	3252	BY SIMILARITY.
FT	DISULFID	3228	3244	BY SIMILARITY.
FT	DISULFID	3259	3302	BY SIMILARITY.
FT	DISULFID	3288	3315	BY SIMILARITY.
FT	CARBOHYD	57	57	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	330	330	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	351	351	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	441	441	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	807	807	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	914	914	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	951	951	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1305	1305	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1372	1372	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1679	1679	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	2054	2054	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	2244	2244	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	2362	2362	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	2627	2627	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	3030	3030	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	3332	3332	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	3342	3342	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	VARSPLIC	348	348	P -> R (in isoform V1 and isoform V3).
FT	VARSPLIC			/FTid=VSP 003087.
FT	VARSPLIC	349	1308	Missing (In isoform V1).
FT	VARSPLIC			/FTid=VSP 003088.
FT	VARSPLIC	1309	3052	Missing (In isoform V2).
FT	VARSPLIC			/FTid=VSP 003089.
FT	VARSPLIC	349	3052	Missing (In isoform V3).
FT	VARSPLIC			/FTid=VSP 003090.
FT	CONFLICT	126	126	A -> G (IN REF. 3).
FT	CONFLICT	348	348	MISSING (IN REF. 3).
FT	CONFLICT	1658	1658	I -> T (IN REF. 3).
FT	CONFLICT	1674	1680	TVWNSNS -> QFGIQTA (IN REF. 3).
SQ	SEQUENCE	3358 AA;	366938 MW;	071B80026BC0762D CRC64;
Query Match 15.6%; Score 174; DB 1; Length 3358;				
Best Local Similarity 28.5%; Pred. No. 5e-07;				
Matches 47; Conservative 23; Mismatches 55; Indels 40; Gaps 8;				
QY	17	QRPCYKVIYFHTSRRLNFEAKEACRRDGGQLVSISEDEQKLEFIENLLPSDGF-	75	
Db	3140	QGQCYK--YF---AHRRTWDAARECRLOGAHLTSILSHEEQMFVNRV-----GH	3187	YQ
QY	76	WIGLRRREKQSNSTACQDLAYATDGSISQFRNWWYVDEP---SCGSEVCVWYHQPSAP	131	
Db	3188	WIGL-----NDKMFHDFRWTGDSALQYENWRPNQPDSPFSFSGEDCVVIWHENG--	3237	
QY	132	AGIGGPMYFOWNDDRCNMKNFICKYS----DEKPAVPSREAEGE	172	
Db	3238	-----QWNDVPCNYHLTYTCKKGTGTVACGQPPVWENAKTFGK	3273	
RESULT 10				
PGCV_HUMAN				
ID	PGCV_HUMAN	STANDARD;	PRT;	3396 AA.
AC	P13611; P20754; Q13010; Q13189; Q15123; Q9UNW5;			
DT	01-JAN-1990 (Rel. 13, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Versican core protein precursor (large fibroblast proteoglycan)			

DE	(Chondroitin sulfate proteoglycan core protein 2) (PG-M) (Glial hyaluronate-binding protein) (GHAP).
DE	CSPG2.
GN	Homo sapiens (Human).
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC	NCBI_TaxID=9606;
OX	[1]
RN	SEQUENCE FROM N.A. (ISOFORM V0).
RP	MEDLINE=95105188; PubMed=7528742;
RX	Naso M.F., Zimmermann D.R., Iozzo R.V.;
RA	"Characterization of the complete genomic structure of the human versican gene and functional analysis of its promoter.";
RT	J. Biol. Chem. 269:32999-33008(1994).
RL	[2]
RN	SEQUENCE FROM N.A. (ISOFORM V1).
RP	TISSUE=Placenta;
RC	MEDLINE=90059882; PubMed=2583089;
RX	Zimmermann D.R., Ruoslahti E.;
RA	"Multiple domains of the large fibroblast proteoglycan, versican.";
RT	EMBO J. 8:2975-2981(1989).
RL	[3]
RN	SEQUENCE FROM N.A. (ISOFORM V2).
RP	TISSUE=Glial tumor;
RC	MEDLINE=95105187; PubMed=7806529;
RX	Dours-Zimmermann M.T., Zimmermann D.R.;
RA	"A novel glycosaminoglycan attachment domain identified in two alternative splice variants of human versican.";
RT	J. Biol. Chem. 269:32992-32998(1994).
RL	[4]
RN	SEQUENCE OF 2711-3396 FROM N.A.
RP	TISSUE=Lung fibroblast;
RC	MEDLINE=88007514; PubMed=2820964;
RX	Krusius T., Gehlsen K.R., Ruoslahti E.;
RA	"A fibroblast chondroitin sulfate proteoglycan core protein contains lectin-like and growth factor-like sequences.";
RT	J. Biol. Chem. 262:13120-13125(1987).
RL	[5]
RN	SEQUENCE OF 251-347 FROM N.A.
RP	MEDLINE=93122792; PubMed=1478664;
RX	Iozzo R.V., Naso M.F., Cannizzaro L.A., Wasmuth J.J., McPherson J.D.;
RA	"Mapping of the versican proteoglycan gene (CSPG2) to the long arm of human chromosome 5 (5q12-5q14).";
RT	Genomics 14:845-851(1992).
RL	[6]
RN	SEQUENCE FROM N.A. (ISOFORM V3).
RP	TISSUE=Brain;
RC	MEDLINE=95181355; PubMed=7876137;
RX	Zako M., Shinomura T., Ujita M., Ito K., Kimata K.;
RA	"Expression of PG-M(V3), an alternatively spliced form of PG-M without a chondroitin sulfate attachment in region in mouse and human tissues.";
RT	J. Biol. Chem. 270:3914-3918(1995).
RL	[7]
RN	SEQUENCE OF 3333-3396 FROM N.A. (ISOFORM VINT).
RP	TISSUE=Aortic smooth muscle;
RC	MEDLINE=99327053; PubMed=10397680;
RX	Lemire J.M., Braun K.R., Maurel P., Kaplan E.D., Schwartz S.M., Wight T.N.;
RA	"Versican/PG-M isoforms in vascular smooth muscle cells.";
RT	Arterioscler. Thromb. Vasc. Biol. 19:1630-1639(1999).
RL	[8]
RN	PARTIAL SEQUENCE.
RP	TISSUE=Brain;
RC	MEDLINE=89174663; PubMed=2466833;
RX	Perides G., Lane W.S., Andrews D., Dahl D., Bignami A.;
RA	"Isolation and partial characterization of a glial hyaluronate-binding protein.";
RT	J. Biol. Chem. 264:5981-5987(1989).
RL	[9]
RN	TISSUE SPECIFICITY OF ISOFORMS.
RP	MEDLINE=96213482; PubMed=8627343;

RA Paulus W., Baur I., Dours-Zimmermann M.T., Zimmermann D.R.;
RT "Differential expression of versican isoforms in brain tumors."
RL J. Neuropathol. Exp. Neurol. 55:528-533(1996).
CC -!- FUNCTION: May play a role in intracellular signaling and in
CC connecting cells with the extracellular matrix. May take part in
CC the regulation of cell motility, growth and differentiation. Binds
CC hyaluronic acid.
CC -!- SUBUNIT: Interacts with FBLN1 (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=5;
CC Comment=Additional isoforms seem to exist;
CC Name=V0;
CC IsoId=P13611-1; Sequence=Displayed;
CC Name=V1;
CC IsoId=P13611-2; Sequence=VSP_003082, VSP_003083;
CC Name=V2;
CC IsoId=P13611-3; Sequence=VSP_003084;
CC Name=V3;
CC IsoId=P13611-4; Sequence=VSP_003082, VSP_003085;
CC Name=Vint;
CC IsoId=P13611-5; Sequence=VSP_003086;
CC -!- TISSUE SPECIFICITY: Cerebral white matter. V0 and V1 is expressed
CC in normal brain, gliomas, medulloblastomas, schwannomas,
CC neurofibromas, and meningiomas; v2 is restricted to normal brain
CC and gliomas; v3 is found in all these tissues except
CC medulloblastomas.
CC -!- DEVELOPMENTAL STAGE: Disappears after the cartilage development.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -!- SIMILARITY: Contains 2 link domains.
CC -!- SIMILARITY: Contains 2 EGF-like domains.
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC -!- SIMILARITY: Contains 1 Sushi (SCR) domain.
CC -!- SIMILARITY: BELONGS TO THE AGGRECAN/VERSICAN PROTEOGLYCAN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U16306; AAA65018.1; -
CC EMBL; X15998; CAA34128.1; -
CC EMBL; S52488; AAB24878.1; -
CC EMBL; U26555; AAA67565.1; -
CC EMBL; D32039; BAA06801.1; -
CC EMBL; J02814; AAA36437.1; -
CC EMBL; AF084545; AAD48545.1; -
CC PIR; S06014; A60979.
CC HSSP; P01132; 1EGF.
CC Genew; HGNC:2464; CSPG2.
CC MIM; 118661; -
CC GO; GO:0005578; C:extracellular matrix; TAS.
CC GO; GO:0005204; F:chondroitin sulfate proteoglycan; TAS.
CC GO; GO:0005540; F:hyaluronic acid binding activity; TAS.
CC GO; GO:0008037; P:cell recognition; TAS.
CC GO; GO:0007275; P:development; TAS.
CC InterPro; IPR000152; Asx hydroxyl.
CC InterPro; IPR000742; EGF_2.
CC InterPro; IPR001881; EGF_Ca.
CC InterPro; IPR006209; EGF_like.
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR003599; Ig.
CC InterPro; IPR003006; Ig_MHC.
CC InterPro; IPR001304; Lectin_C.
CC InterPro; IPR000538; Link.
CC InterPro; IPR000436; Sushi_SCR_CCP.
CC Pfam; PF00008; EGF; 2.
CC Pfam; PF00059; lectin_c; 1.
CC Pfam; PF00084; sushi; 1.
CC Pfam; PF00193; Xlink; 2.

DR ProDom; PD000918; Link; 2.
DR SMART; SM00032; CCP; 1.
DR SMART; SM00034; CLECT; 1.
DR SMART; SM00179; EGF_CA; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00445; LINK; 2.
DR PROSITE; PS00010; ASX_HYDROXYL; 1.
DR PROSITE; PS00615; C_TYPE_LLECTIN_1; 1.
DR PROSITE; PS00041; C_TYPE_LLECTIN_2; 1.
DR PROSITE; PS00022; EGF_1_2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01187; EGF_CA; 1.
DR PROSITE; PS00835; IG_LIKE; 1.
DR PROSITE; PS01241; LINK; 2.
KW Glycoprotein; Proteoglycan; Lectin; Extracellular matrix; Sushi;
KW Signal; Repeat; EGF-like domain; Calcium; Immunoglobulin domain;
KW Hyaluronic acid; Alternative splicing.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 3396 VERSICAN CORE PROTEIN.
FT DOMAIN 21 146 IG-LIKE V-TYPE.
FT DOMAIN 167 244 LINK 1.
FT DOMAIN 265 346 LINK 2.
FT DOMAIN 348 1335 GAG-ALPHA
(GLUCOSAMINOGLYCAN ATTACHMENT DOMAIN).
FT DOMAIN 1336 3089 GAG-BETA.
FT DOMAIN 3089 3125 EGF-LIKE 1.
FT DOMAIN 3127 3163 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 3176 3290 C-TYPE LECTIN.
FT DOMAIN 3295 3353 SUSHI.
FT DISULFID 44 130 BY SIMILARITY.
FT DISULFID 172 243 BY SIMILARITY.
FT DISULFID 196 217 BY SIMILARITY.
FT DISULFID 270 345 BY SIMILARITY.
FT DISULFID 294 315 BY SIMILARITY.
FT DISULFID 3093 3104 BY SIMILARITY.
FT DISULFID 3098 3113 BY SIMILARITY.
FT DISULFID 3115 3124 BY SIMILARITY.
FT DISULFID 3131 3142 BY SIMILARITY.
FT DISULFID 3136 3151 BY SIMILARITY.
FT DISULFID 3153 3162 BY SIMILARITY.
FT DISULFID 3169 3180 BY SIMILARITY.
FT DISULFID 3197 3289 BY SIMILARITY.
FT DISULFID 3265 3281 BY SIMILARITY.
FT DISULFID 3296 3339 BY SIMILARITY.
FT DISULFID 3325 3352 BY SIMILARITY.
FT CARBOHYD 57 57 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 330 330 N-LINKED (GLCNAC. .).
FT CARBOHYD 615 615 N-LINKED (GLCNAC. .) (POTENTIAL).

Query Match 15.6%; Score 174; DB 1; Length 3396;
Best Local Similarity 28.5%; Pred. No. 5.1e-07;
Matches 47; Conservative 23; Mismatches 55; Indels 40; Gaps 8;

QY 17 QRPCYKVIYFHDTSRRLNFEAEKACRRDGGQLVSISESEDEQKLIKFIENLLPSDGF- 75
Db 3177 QGQCYK--YF---AHRRTWDAARECRLOGAHLTSLSHEEQMFVNRV-----GHDYQ 3224

QY 76 WIGLRRREEKQSNSTACQDLYAWTDGSIQFRNRYVDEP-----SCGSEVCVVMYHQPSAP 131
Db 3225 WIGL-----NDKMFHDHFRWTDGSTLQYENWRPNQDPSFFSAGEDCVVIWHENG-- 3274

QY 132 AGIGGPMFQWDDRCNMKNFICKYS-----DEKPAVPSREAGE 172
Db 3275 -----QWNVPCNYHLTYTCKKGTVACGQPPVVENAKTFGK 3310

RESULT 11
PGCV-CHICK
ID -PGCV CHICK STANDARD; PRT; 3562 AA.
AC Q90953; Q90945;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Versican core protein precursor (large fibroblast proteoglycan)
DE (Chondroitin sulfate proteoglycan core protein 2) (PG-M).
GN CPSG2.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS V0 AND V1).
RC STRAIN=White leghorn; TISSUE=Limb bud;
RX MEDLINE=93300846; PubMed=8314802;
RA Shinomura T., Nishida Y., Ito K., Kimata K.;
RT "cDNA cloning of PG-M, a large chondroitin sulfate proteoglycan
RT expressed during chondrogenesis in chick limb buds. Alternative
RT spliced multiforms of PG-M and their relationships to versican.";
RL J. Biol. Chem. 268:14461-14469(1993).
CC -!- FUNCTION: May play a role in intercellular signaling and in
CC connecting cells with the extracellular matrix. May take part in
CC the regulation of cell motility, growth and differentiation. Binds
CC hyaluronic acid.
CC -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Comment=Additional isoforms seem to exist;
CC Name=V0;
CC IsoId=Q90953-1; Sequence=Displayed;
CC Name=V1;
CC IsoId=Q90953-2; Sequence=VSP_003093;
CC TISSUE SPECIFICITY: Prechondrogenic condensation area of
CC developing limb buds.
CC -!- DEVELOPMENTAL STAGE: Disappears after the cartilage development
CC (by similarity).
CC -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -!- SIMILARITY: Contains 2 link domains.
CC -!- SIMILARITY: Contains 2 EGF-like domains.
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC -!- SIMILARITY: Contains 1 Sushi (SCR) domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X60226; CAA42787.1; -.
DR EMBL; D13542; BAA02742.1; -.
DR PIR; A47171; A47171.
DR HSSP; P00740; 1EDM.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR007110; Ig_Like.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR000538; Link.
DR InterPro; IPR000436; Sushi_SCR_CCP.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF00059; lectin_c; 1.
DR Pfam; PF00084; sushi; 1.
DR Pfam; PF00193; Xlink; 2.
DR PRINTS; PR01265; LINKMODULE.
DR ProDom; PD000918; Link; 2.
DR SMART; SM00032; CCP; 1.
DR SMART; SM00034; CLECT; 1.
DR SMART; SM00179; EGF_CA; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00445; LINK; 2.

DR PROSITE; PS00010; ASX_HYDROXYL; 1.
DR PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR PROSITE; PS00041; C_TYPE_LECTIN_2; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01187; EGF_CA; 1.
DR PROSITE; PS00835; IG_LIKE; 1.
DR PROSITE; PS01241; LINK; 2.
KW Glycoprotein; Proteoglycan; Lectin; Extracellular matrix; Sushi;
KW Signal; Repeat; EGF-like domain; Calcium; Immunoglobulin domain;
KW Hyaluronic acid; Alternative splicing.
FT SIGNAL 1 26
FT CHAIN 27 3562
FT DOMAIN 27 143
FT DOMAIN 166 243
FT DOMAIN 264 345
FT DOMAIN 3254 3290
FT DOMAIN 3292 3328
FT DOMAIN 3341 3455
FT DOMAIN 3460 3518
FT DISULFID 44 129
FT DISULFID 171 242
FT DISULFID 195 216
FT DISULFID 269 344
FT DISULFID 293 314
FT DISULFID 3258 3269
FT DISULFID 3263 3278
FT DISULFID 3280 3289
FT DISULFID 3296 3307
FT DISULFID 3301 3316
FT DISULFID 3318 3327
FT DISULFID 3334 3345
FT DISULFID 3362 3454
FT DISULFID 3430 3446
FT DISULFID 3461 3504
FT DISULFID 3490 3517
FT CARBOHYD 163 163
FT CARBOHYD 235 235
FT CARBOHYD 329 329
FT CARBOHYD 529 529
FT CARBOHYD 709 709
FT CARBOHYD 948 948
FT CARBOHYD 1409 1409
FT CARBOHYD 1479 1479
FT CARBOHYD 1523 1523
FT CARBOHYD 1530 1530
FT CARBOHYD 1625 1625
FT CARBOHYD 1751 1751
FT CARBOHYD 1988 1988
FT CARBOHYD 2088 2088
FT CARBOHYD 2089 2089
FT CARBOHYD 2507 2507
FT CARBOHYD 2642 2642
FT CARBOHYD 2679 2679
FT CARBOHYD 2748 2748
FT CARBOHYD 2762 2762
FT CARBOHYD 3069 3069
FT CARBOHYD 3194 3194
FT CARBOHYD 3232 3232
FT CARBOHYD 3545 3545
FT VARSPIC 485 1411
SQ SEQUENCE 3562 AA; 388078 MW; 9BC566E98C1602D2 CRC64;
/FTId=VSP_003093.

Query Match 15.3%; Score 171; DB 1; Length 3562;
Best Local Similarity 28.5%; Pred. No. 9.9e-07;
Matches 47; Conservative 23; Mismatches 55; Indels 40; Gaps 8;

QY 17 QRPCYKVIYFHDTSRRLNFEFEAKEACRRDGGQLVLSIESEDEQKLIKIEFIENLLPSDGF- 75
Db 3342 QGQCYK--YF---AHRRTWDTAERECLQGAHLTSILSHEEQVFNRI-----GHDIQ 3389
QY 76 WIGLRREEKQSNSTACQDLYAWTDGSIQSFRNWWYVDEP-----SCGSEVCVVMYHQSAP 131

Db 3390 WIGL-----NDKMFDFRWDGSPLOQYENWRPNQDSFFSAGEDCVVIWHENG-- 3439

QY 132 AGIGSPYMFQWDDRCNMKNFICKYS----DEKPAVPSREAEGE 172

Db 3440 -----QWNDVPCNYHLTYTCKKGTACGQPPVVENAKTFGK 3475

RESULT 12

CD93_RAT

ID CD93_RAT STANDARD; PRT; 643 AA.

AC Q9ET61; Q9JIZ6;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Complement component C1q receptor precursor (Complement component 1, q

DE subcomponent, receptor 1) (C1qRp) (C1qR(p)) (C1q/MBL/SPA receptor)

DE (CD93 antigen) (Cell surface antigen AA4).

GN C1QR1 OR CD93 OR C1QRP.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=PVG; TISSUE=Natural killer cells;

RX MEDLINE=20545218; PubMed=11093152;

RA Lovik G., Vaage J.T., Dissen E., Szpirer C., Ryan J.C., Rolstad B.;

RT "Characterization and molecular cloning of rat C1qRp, a receptor on NK

RT cells.";

RL Eur. J. Immunol. 30:3355-3362(2000).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=Wistar; TISSUE=Lung;

RX MEDLINE=20507883; PubMed=10934210;

RA Dean Y.D., McGreal E.P., Akatsu H., Gasque P.;

RT "Molecular and cellular properties of the rat AA4 antigen, a C-type

RT lectin-like receptor with structural homology to thrombomodulin.";

RL J. Biol. Chem. 275:34382-34392(2000).

CC -!- FUNCTION: Receptor (or element of a larger receptor complex) for

CC Clq, mannose-binding lectin (MBL2) and pulmonary surfactant

CC protein A (SPA). May mediate the enhancement of phagocytosis in

CC monocytes and macrophages upon interaction with soluble defense

CC collagens. May play a role in intercellular adhesion.

CC -!- SUBCELLULAR LOCATION: Type I membrane protein.

CC -!- TISSUE SPECIFICITY: Widely expressed. Highly expressed in lung and

CC heart. Expressed at lower level in brain, thymus, liver, spleen,

CC intestine, kidney, adrenal gland, muscle and testis. Expressed on

CC endothelial cells, platelets, undifferentiated monocytes and

CC circulating natural killer cells.

CC -!- PTM: N- and O-glycosylated (By similarity).

CC -!- SIMILARITY: Contains 1 C-type lectin family domain.

CC -!- SIMILARITY: Contains 5 EGF-like domains.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL; AF136537; AAG01572.1; -.

CC EMBL; AF160978; AAF80402.1; -.

CC HSSP; P35555; 1EMN.

DR GO; GO:0016021; C:integral to membrane; ISS.

DR GO; GO:0004872; F:receptor activity; ISS.

DR GO; GO:0016337; P:cell-cell adhesion; ISS.

DR GO; GO:0042116; P:macrophage activation; ISS.

DR GO; GO:0006909; P:phagocytosis; ISS.

DR InterPro; IPR00152; Asx_hydroxyl.

DR InterPro; IPR001881; EGF_Ca.

DR InterPro; IPR006209; EGF_like.

DR InterPro; IPR001304; Lectin_C.

DR Pfam; PF00008; EGF; 5.

DR Pfam; PF00059; lectin_c; 1.

DR SMART; SM00034; CLECT; 1.

DR SMART; SM00179; EGF_CA; 3.

DR PROSITE; PS00010; ASX_HYDROXYL; 3.

DR PROSITE; PS00615; C_TYPE_LECTIN_1; FALSE_NEG.

DR PROSITE; PS00041; C_TYPE_LECTIN_2; 1.

DR PROSITE; PS01186; EGF_2; 3.

DR PROSITE; PS01187; EGF_CA; 3.

KW Cell adhesion; Receptor; Repeat; Signal; Transmembrane;

KW EGF-like domain; Lectin; Glycoprotein.

-FT SIGNAL 1 23 POTENTIAL.

FT CHAIN 24 643 COMPLEMENT COMPONENT C1Q RECEPTOR.

FT DOMAIN 24 571 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 572 592 POTENTIAL.

FT DOMAIN 593 643 CYTOPLASMIC (POTENTIAL).

FT DOMAIN 31 173 C-TYPE LECTIN.

FT DOMAIN 257 298 EGF-LIKE 1.

FT DOMAIN 299 341 EGF-LIKE 2.

FT DOMAIN 342 381 EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).

FT DOMAIN 382 423 EGF-LIKE 4, CALCIUM-BINDING (POTENTIAL).

FT DOMAIN 424 462 EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).

FT DISULFID 261 272 BY SIMILARITY.

FT DISULFID 268 282 BY SIMILARITY.

FT DISULFID 284 297 BY SIMILARITY.

FT DISULFID 303 314 BY SIMILARITY.

FT DISULFID 308 325 BY SIMILARITY.

FT DISULFID 327 340 BY SIMILARITY.

FT DISULFID 346 355 BY SIMILARITY.

FT DISULFID 351 364 BY SIMILARITY.

FT DISULFID 366 380 BY SIMILARITY.

FT DISULFID 386 397 BY SIMILARITY.

FT DISULFID 393 406 BY SIMILARITY.

FT DISULFID 408 422 BY SIMILARITY.

FT DISULFID 428 437 BY SIMILARITY.

FT DISULFID 433 446 BY SIMILARITY.

FT DISULFID 448 461 BY SIMILARITY.

FT CARBOHYD 322 322 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 498 498 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CONFLICT 417 417 E -> K (IN REF. 2).

SQ SEQUENCE 643 AA; 68781 MW; 9AE4C933AD943DB6 CRC64;

Query Match 14.8%; Score 165; DB 1; Length 643;

Best Local Similarity 25.2%; Pred. No. 4.2e-07;

Matches 52; Conservative 35; Mismatches 73; Indels 46; Gaps 10;

Qy 3 GRLLSGQP-----VCRGGTQPCYKVIYFHDTSRRLNFEAEACRRDGGQLVSI ESD 56

Db 15 GQLWAGAAADSEAVVCEG---TACYTAHW-----GKLSAAEAQHRCNENGGNLATVKSEE 66

Qy 57 EQKLEKPIENLL-----PSD---GDFWIGLRRREEKQSNSTACQDLYAWT-DGSISQFRN 108

Db 67 EARHVQEALAQLLKTKAPSETKICKFWIGLQREKCKTYHDLPMKGFSWVGGGDTTYSN 126

Qy 109 WY-VDEPSCGSEVCVVMY-----HQPSAPAGIGGYPYMFQWND DRC-----NMKNNF 153

Db 127 WYKASKSSCISKRCVSLILDLSLKPSPHLP-----KWHESPCGTPDAPGNSIEGF 177

Qy 154 IKYSDEKPAVPSREAEGEETELTTP 179

Db 178 LCKFNFKGMCSPALGPGQLTYTTP 203

RESULT 13

PGCA_BOVIN

ID PGCA_BOVIN STANDARD; PRT; 2364 AA.

AC P13608; P79117; Q28159;

DT 01-JAN-1990 (Rel. 13, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Aggreacan core protein precursor (Cartilage-specific proteoglycan core

DE protein) (CSPCP).

GN AGC1.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Hering T.M., Kollar J., Huynh T.D.;
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 563-1056 FROM N.A.
RX MEDLINE=89380219; PubMed=2528543;
RA Antonsson P., Heinegaard D., Oldberg A.;
RT "The keratan sulfate-enriched region of bovine cartilage proteoglycan
consists of a consecutively repeated hexapeptide motif.";
RL J. Biol. Chem. 264:16170-16173(1989).
RN [3]
RP SEQUENCE OF 1609-2113 AND 2151-2364 FROM N.A.
RX MEDLINE=87270630; PubMed=3111460;
RA Oldberg A., Antonsson P., Heinegaard D.;
RT "The partial amino acid sequence of bovine cartilage proteoglycan,
deduced from a cDNA clone, contains numerous Ser-Gly sequences
arranged in homologous repeats.";
RL Biochem. J. 243:255-259(1987).
RN [4]
RP SEQUENCE OF 2114-2150 FROM N.A.
RC TISSUE=Cartilage;
RX MEDLINE=93352525; PubMed=8349621;
RA Fueleop C., Walcz E., Vallyon M., Glant T.T.;
RT "Expression of alternatively spliced epidermal growth factor-like
domains in aggrecans of different species. Evidence for a novel
module.";
RL J. Biol. Chem. 268:17377-17383(1993).
RN [5]
RP PARTIAL SEQUENCE.
RX MEDLINE=85027710; PubMed=6489519;
RA Perin J.-P., Bonnet F., Jolles J., Jolles P.;
RT "Sequence data concerning the protein core of the cartilage
proteoglycan monomers. Characterization of a sequence allowing the
synthesis of an oligonucleotide probe.";
RL FEBS Lett. 176:37-42(1984).
RN [6]
RP PARTIAL SEQUENCE.
RX MEDLINE=87005253; PubMed=3530809;
RA Perin J.P., Bonnet F., Jolles P.;
RT "Structural relationship between link proteins and proteoglycan
monomers.";
RL FEBS Lett. 206:73-77(1986).
CC -!- FUNCTION: THIS PROTEOGLYCAN IS A MAJOR COMPONENT OF EXTRACELLULAR
MATRIX OF CARTILAGINOUS TISSUES. A MAJOR FUNCTION OF THIS PROTEIN
IS TO RESIST COMPRESSION IN CARTILAGE. IT BINDS AVIDLY TO
HYALURONIC ACID VIA AN AMINO-TERMINAL GLOBULAR REGION. MAY PLAY A
REGULATORY ROLE IN THE MATRIX ASSEMBLY OF THE CARTILAGE.
CC -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix (By
similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=P13608-1; Sequence=Displayed;
CC Name=2;
CC IsoId=P13608-2; Sequence=VSP_003072;
CC -!- DOMAIN: TWO GLOBULAR DOMAINS, G1 AND G2, COMPRISE THE AMINO
TERMINUS OF THE PROTEOGLYCAN, WHILE ANOTHER GLOBULAR REGION, G3,
MAKES UP THE C-TERMINUS. G1 CONTAINS LINK DOMAINS AND THUS
CONSISTS OF THREE DISULFIDE-BONDED LOOP STRUCTURES DESIGNATED AS
THE A, B, B' MOTIFS. G2 IS SIMILAR TO G1. THE KERATAN SULFATE (KS)
AND THE CHONDROITIN SULFATE (CS) ATTACHMENT DOMAINS LIE BETWEEN G2
AND G3.
CC -!- PTM: CONTAINS MOSTLY CHONDROITIN SULFATE, BUT ALSO N-LINKED AND O-
LINKED (ABOUT 40) OLIGOSACCHARIDES.
CC -!- PTM: THE KERATAN SULFATE CONTENTS DIFFER CONSIDERABLY BETWEEN
ADULT AND FETAL BOVINE PROTEOGLYCANS.

CC -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -!- SIMILARITY: Contains 4 link domains.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC -!- SIMILARITY: Contains 1 Sushi (SCR) domain.
CC -!- SIMILARITY: BELONGS TO THE AGGREGAN/VERSICAN PROTEOGLYCAN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U76615; AAB38524.1; -;
DR EMBL; L07053; -; NOT_ANNOTATED_CDS.
DR PIR; A34234; A39808.
DR PIR; T42630; T42630.
DR HSSP; P08709; 1BF9.
DR InterPro; IPR002353; AntifreezeII.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR007110; Ig_Like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR000538; Link.
DR InterPro; IPR003324; SGXSG.
DR InterPro; IPR000436; Sushi_SCR_CCP.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00059; lectin_c; 1.
DR Pfam; PF02339; SGXSG; 61.
DR Pfam; PF00084; sushi; 1.
DR Pfam; PF00193; Xlink; 4.
DR PRINTS; PR01265; LINKMODULE.
DR PRINTS; PR00356; ANTIFREEZEII.
DR ProDom; PD000918; Link; 4.
DR SMART; SM00032; CCP; 1.
DR SMART; SM00034; CLECT; 1.
DR SMART; SM00179; EGF_CA; 1.
DR SMART; SM00445; LINK; 4.
DR PROSITE; PS00010; ASX_HYDROXYL; 1.
DR PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01187; EGF_CA; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR PROSITE; PS00290; IG_MHC; FALSE_NEG.
DR PROSITE; PS01241; LINK; 4.
KW Glycoprotein; Cartilage; Proteoglycan; Lectin; Signal; Sushi;
KW EGF-like domain; Calcium; Alternative splicing; Repeat;
KW Immunoglobulin domain.
FT SIGNAL 1 16 POTENTIAL.
FT CHAIN 17 2364 AGGREGAN CORE PROTEIN.
FT DOMAIN 25 147 IG-LIKE V-TYPE.
FT DOMAIN 170 247 LINK 1.
FT DOMAIN 268 349 LINK 2.
FT DOMAIN 504 581 LINK 3.
FT DOMAIN 602 683 LINK 4.
FT DOMAIN 774 907 23 X 6 AA APPROXIMATE TANDEM REPEATS OF
E-[EK]-P-F-P-S.
FT CS-2.
FT DOMAIN 1433 2112 EGF-LIKE, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 2113 2149 G3.
FT DOMAIN 2114 2364 C-TYPE LECTIN.
FT DOMAIN 2161 2276 SUSHI.
FT DOMAIN 2280 2338 BY SIMILARITY.
FT DISULFID 51 133 BY SIMILARITY.
FT DISULFID 175 246 BY SIMILARITY.
FT DISULFID 199 220 BY SIMILARITY.
FT DISULFID 273 348 BY SIMILARITY.

DR PRINTS; PR01265; LINKMODULE.
DR PRINTS; PR00356; ANTIFREEZEII.
DR ProDom; PD000918; Link; 4.
DR SMART; SM00032; CCP; 1.
DR SMART; SM00034; CLECT; 1.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00445; LINK; 4.
DR PROSITE; PS00615; C-TYPE LECTIN_1; 1.
DR PROSITE; PS50041; C-TYPE LECTIN_2; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR PROSITE; PS00290; IG_MHC; 1.
DR PROSITE; PS01241; LINK; 3.
KW Glycoprotein; Cartilage; Proteoglycan; Lectin; Signal; Sushi;
KW EGF-like domain; Alternative potential; Repeat; Immunoglobulin domain.
FT SIGNAL 1 19
FT CHAIN 20 2415
FT DOMAIN 34 147
FT DOMAIN 170 247
FT DOMAIN 268 349
FT DOMAIN 495 572
FT DOMAIN 593 673
FT DOMAIN 2164 2199
FT DOMAIN 2201 2327
FT DOMAIN 2331 2389
FT DOMAIN 48 141
FT DOMAIN 152 247
FT DOMAIN 253 349
FT DOMAIN 477 571
FT DOMAIN 578 672
FT DOMAIN 676 848
FT DOMAIN 772 844
FT DOMAIN 851 1497
FT DOMAIN 941 1497
FT DOMAIN 1498 2162
FT DOMAIN 2163 2415
FT DISULFID 51 133
FT DISULFID 175 246
FT DISULFID 199 220
FT DISULFID 273 348
FT DISULFID 297 318
FT DISULFID 500 571
FT DISULFID 524 545
FT DISULFID 598 672
FT DISULFID 621 642
FT DISULFID 2168 2178
FT DISULFID 2173 2187
FT DISULFID 2189 2198
FT DISULFID 2205 2216
FT DISULFID 2233 2325
FT DISULFID 2301 2317
FT DISULFID 2332 2375
FT DISULFID 2361 2388
FT CARBOHYD 126 126
FT CARBOHYD 239 239
FT CARBOHYD 333 333
FT CARBOHYD 387 387
FT CARBOHYD 434 434
FT CARBOHYD 602 602
FT CARBOHYD 657 657
FT CARBOHYD 737 737
FT CARBOHYD 1898 1898
FT VARSPLIC 2163 2200
FT VARSPLIC 2330 2390
FT CONFLICT 766 766
FT CONFLICT 847 847
FT CONFLICT 1928 1928
FT CONFLICT 1964 1964
FT CONFLICT 2070 2070

FT CONFLICT 2391 2391 A -> P (IN REF. 2 AND 3).
SQ SEQUENCE 2415 AA; 250191 MW; 1288937E1B98C6B6 CRC64;
Query Match 14.2%; Score 158.5; DB 1; Length 2415;
Best Local Similarity 29.1%; Pred. No. 8.1e-06;
Matches 52; Conservative 22; Mismatches 62; Indels 43; Gaps 11;
QY 9 QPVCRRGG---TORPCYKVIYFHDTSRRLLNFEEAKEACRRDGGQLVSISEDEQKLIKFI 65
Db 2202 QEVCEEGWNKYQGHYR--HFPD---RETWVDAERRCRREQQSHLSSIVTPEEQ----EPV 2252
QY 66 ENLLPSDGDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNWWYVDEPS---CGSEVC 121
Db 2253 NN---NAQDYQWIGL-----NDRTIEGDFRWSGDGHPMQFENWRPNQDNFFAAGEDC 2301
QY 122 VVM-YHQPSAPAGIGGYPMFQWDDRCNMKNFNICKYS----DEKPAVPSREAEGETE 175
Db 2302 VVMIWHEKG-----EWNDVPCNYHLPTCKKGTACGEPVVEHARTFGQKKD 2349
RESULT 15
PGCA_CANFA
ID PGCA_CANFA STANDARD; PRT; 2333 AA.
AC Q28343; Q28310;
DT 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Aggrecan core protein precursor (Cartilage-specific proteoglycan core protein) (CSPCP).
DE AGC1.
GN Canis familiaris (Dog).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RA Glant T.T., Adams M.E., Kwok S.X.F., Huang D., Fueleop C.;
RT "Complete coding sequence and deduced amino acid sequence of aggrecan of canine cartilage."
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 774-833 FROM N.A.
RC TISSUE=Cartilage;
RX MEDLINE=95128522; PubMed=7827755;
RA Barry F.P., Neame P.J., Sasse J., Pearson D.;
RT "Length variation in the keratan sulfate domain of mammalian aggrecan."
RL Matrix Biol. 14:323-328(1994).
RN [3]
RP SEQUENCE OF 1830-2333 FROM N.A.
RA Adams M.E., Kwok S.X.F., Huang D., Glant T.T., Fullop C.;
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 2082-2118 FROM N.A.
RC TISSUE=Cartilage;
RX MEDLINE=93352525; PubMed=8349621;
RA Fueleop C., Walcz E., Valyon M., Glant T.T.;
RT "Expression of alternatively spliced epidermal growth factor-like domains in aggrecans of different species. Evidence for a novel module."
RL J. Biol. Chem. 268:17377-17383(1993).
CC -!- FUNCTION: THIS PROTEOGLYCAN IS A MAJOR COMPONENT OF EXTRACELLULAR MATRIX OF CARTILAGINOUS TISSUES. A MAJOR FUNCTION OF THIS PROTEIN IS TO RESIST COMPRESSION IN CARTILAGE. IT BINDS AVIDLY TO HYALURONIC ACID VIA AN AMINO-TERMINAL GLOBULAR REGION. MAY PLAY A REGULATORY ROLE IN THE MATRIX ASSEMBLY OF THE CARTILAGE.
CC -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix (By similarity).
CC -!- DOMAIN: TWO GLOBULAR DOMAINS, G1 AND G2, COMPRISE THE AMINO TERMINUS OF THE PROTEOGLYCAN, WHILE ANOTHER GLOBULAR REGION, G3, MAKES UP THE COOH TERMINUS. G1 CONTAINS LINK DOMAINS AND THUS CONSISTS OF THREE DISULFIDE-BONDED LOOP STRUCTURES DESIGNATED AS THE A, B, B' MOTIFS. G2 IS SIMILAR TO G1. THE KERATAN SULFATE (KS)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 22, 2003, 16:06:34 ; Search time 10.0151 Seconds
(without alignments)
1978.090 Million cell updates/sec

Title: US-09-887-855-5
Perfect score: 1115
Sequence: 1 ATGRLLSGQPVCRGGTQRPC.....EEDAKKTFKESREAAALNLAY 206

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	182	16.3	1456	1	A36563	mannose receptor p
2	178.5	16.0	1455	1	A48925	mannose receptor p
3	177	15.9	1268	2	S52781	neurocan - mouse
4	174.5	15.7	1643	2	T14274	versican precursor
5	174.5	15.7	3381	2	T42389	versican precursor
6	174	15.6	1257	2	S28764	neurocan precursor
7	174	15.6	2397	1	A55535	versican precursor
8	174	15.6	2409	1	A60979	versican precursor
9	171	15.3	3562	2	A47171	chondroitin sulfat
10	170.5	15.3	1479	2	T42710	mannose receptor,
11	158.5	14.2	1340	2	A39808	proteoglycan core
12	158.5	14.2	2327	2	T42630	aggrecan - bovine
13	158.5	14.2	2415	1	A39086	aggrecan precursor
14	154.5	13.9	612	2	B42755	E-selectin precurs
15	153.5	13.8	2124	2	A28452	proteoglycan core
16	152	13.6	912	2	A54423	brevican precursor
17	149.5	13.4	459	2	T24425	hypothetical prote
18	149	13.4	321	1	LNHUER	IgE Fc receptor II
19	148.5	13.3	330	2	T46256	brevican - human (
20	148.5	13.3	2132	1	A55182	aggrecan precursor
21	148	13.3	253	2	E89130	protein F52E1.2 (i
22	147	13.2	883	2	S57653	brevican precursor
23	146.5	13.1	162	1	LNRC1	lectin BRA3-1 prec
24	146	13.1	2109	1	I50421	aggrecan precursor
25	145.5	13.0	742	2	JC7595	scavenger receptor
26	145	13.0	883	2	S49126	brevican precursor
27	144.5	13.0	173	2	S10548	lectin - barnacle
28	144.5	13.0	372	2	S23936	L-selectin precurs
29	144.5	13.0	404	2	A46274	HIV gp120-binding

30	143.5	12.9	129	2	JC4329	coagulation factor
31	143.5	12.9	372	1	A32375	L-selectin precurs
32	143.5	12.9	463	2	T26655	hypothetical prote
33	142.5	12.8	131	2	JC5058	bitiscetin alpha c
34	142	12.7	1487	2	S48719	phospholipase-A(2)
35	141.5	12.7	331	1	LNMSER	IgE Fc receptor, 1
36	140.5	12.6	162	1	LNRC3	lectin BRA3-2 prec
37	140	12.6	370	2	S22124	L-selectin precurs
38	139	12.5	248	1	LNHUPS	pulmonary surfacta
39	139	12.5	248	1	LNHUP6	pulmonary surfacta
40	139	12.5	248	1	LNHUP1	pulmonary surfacta
41	139	12.5	283	1	LNPHLS	lectin precursor -
42	138.5	12.4	152	2	JC4690	coagulation factor
43	138.5	12.4	202	2	JC4031	tetranectin precur
44	138	12.4	280	2	T29200	hypothetical prote
45	137.5	12.3	309	1	S34198	IgE Fc receptor II

ALIGNMENTS

RESULT 1

A36563

mannose receptor precursor - human

C;Species: Homo sapiens (man)

C;Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text change 10-Sep-1999

C;Accession: A36563; A60926; A44255; B44255; C44255; F44255; G44255; H44255; I44255; J44255; K44255; L44255; M44255; N44255; O44255; P44255; Q44255; R44255; S44255; T44255; U44255; V44255; W44255; X44255; Y44255; Z44255; AA44255; AB44255; AC44255; AD44255; AE44255; AF44255; AG44255; AH44255; AI44255; AJ44255; AK44255; AL44255; AM44255; AN44255; AO44255; AP44255; AQ44255; AR44255; AS44255; AT44255; AU44255; AV44255; AW44255; AX44255; AY44255; AZ44255; BA44255; BB44255; BC44255; BD44255; BE44255; BF44255; BG44255; BH44255; BI44255; BJ44255; BK44255; BL44255; BM44255; BN44255; BO44255; BP44255; BQ44255; BR44255; BS44255; BT44255; BU44255; BV44255; BW44255; BX44255; BY44255; BZ44255; CA44255; CB44255; CC44255; CD44255; CE44255; CF44255; CG44255; CH44255; CI44255; CJ44255; CK44255; CL44255; CM44255; CN44255; CO44255; CP44255; CQ44255; CR44255; CS44255; CT44255; CU44255; CV44255; CW44255; CX44255; CY44255; CZ44255; DA44255; DB44255; DC44255; DD44255; DE44255; DF44255; DG44255; DH44255; DI44255; DJ44255; DK44255; DL44255; DM44255; DN44255; DO44255; DP44255; DQ44255; DR44255; DS44255; DT44255; DU44255; DV44255; DW44255; DX44255; DY44255; DZ44255; EA44255; EB44255; EC44255; ED44255; EE44255; EF44255; EG44255; EH44255; EI44255; EJ44255; EK44255; EL44255; EM44255; EN44255; EO44255; EP44255; EQ44255; ER44255; ES44255; ET44255; EU44255; EV44255; EW44255; EX44255; EY44255; EZ44255; FA44255; FB44255; FC44255; FD44255; FE44255; FF44255; FG44255; FH44255; FI44255; FJ44255; FK44255; FL44255; FM44255; FN44255; FO44255; FP44255; FQ44255; FR44255; FS44255; FT44255; FU44255; FV44255; FW44255; FX44255; FY44255; FZ44255; GA44255; GB44255; GC44255; GD44255; GE44255; GF44255; GH44255; GI44255; GJ44255; GK44255; GL44255; GM44255; GN44255; GO44255; GP44255; GQ44255; GR44255; GS44255; GT44255; GU44255; GV44255; GW44255; GX44255; GY44255; GZ44255; HA44255; HB44255; HC44255; HD44255; HE44255; HF44255; HG44255; HH44255; HI44255; HJ44255; HK44255; HL44255; HM44255; HN44255; HO44255; HP44255; HQ44255; HR44255; HS44255; HT44255; HU44255; HV44255; HW44255; HX44255; HY44255; HZ44255; IA44255; IB44255; IC44255; ID44255; IE44255; IF44255; IG44255; IH44255; II44255; IJ44255; IK44255; IL44255; IM44255; IN44255; IO44255; IP44255; IQ44255; IR44255; IS44255; IT44255; IU44255; IV44255; IW44255; IX44255; IY44255; IZ44255; JA44255; JB44255; JC44255; JD44255; JE44255; JF44255; JG44255; JH44255; JI44255; JJ44255; JK44255; JL44255; JM44255; JN44255; JO44255; JP44255; JQ44255; JR44255; JS44255; JT44255; JU44255; JV44255; JW44255; JX44255; JY44255; JZ44255; KA44255; KB44255; KC44255; KD44255; KE44255; KF44255; KG44255; KH44255; KI44255; KJ44255; KK44255; KL44255; KM44255; KN44255; KO44255; KP44255; KQ44255; KR44255; KS44255; KT44255; KU44255; KV44255; KW44255; KX44255; KY44255; KZ44255; LA44255; LB44255; LC44255; LD44255; LE44255; LF44255; LG44255; LH44255; LI44255; LJ44255; LK44255; LL44255; LM44255; LN44255; LO44255; LP44255; LQ44255; LR44255; LS44255; LT44255; LU44255; LV44255; LW44255; LX44255; LY44255; LZ44255; MA44255; MB44255; MC44255; MD44255; ME44255; MF44255; MG44255; MH44255; MI44255; MJ44255; MK44255; ML44255; MN44255; MO44255; MP44255; MQ44255; MR44255; MS44255; MT44255; MU44255; MV44255; MW44255; MX44255; MY44255; MZ44255; NA44255; NB44255; NC44255; ND44255; NE44255; NF44255; NG44255; NH44255; NI44255; NJ44255; NK44255; NL44255; NM44255; NO44255; NP44255; NQ44255; NR44255; NS44255; NT44255; NU44255; NV44255; NW44255; NX44255; NY44255; NZ44255; OA44255; OB44255; OC44255; OD44255; OE44255; OF44255; OG44255; OH44255; OI44255; OJ44255; OK44255; OL44255; OM44255; ON44255; OO44255; OP44255; OQ44255; OR44255; OS44255; OT44255; OU44255; OV44255; OW44255; OX44255; OY44255; OZ44255; PA44255; PB44255; PC44255; PD44255; PE44255; PF44255; PG44255; PH44255; PI44255; PJ44255; PK44255; PL44255; PM44255; PN44255; PO44255; PP44255; PQ44255; PR44255; PS44255; PT44255; PU44255; PV44255; PW44255; PX44255; PY44255; PZ44255; QA44255; QB44255; QC44255; QD44255; QE44255; QF44255; QG44255; QH44255; QI44255; QJ44255; QK44255; QL44255; QM44255; QN44255; QO44255; QP44255; QQ44255; QR44255; QS44255; QT44255; QU44255; QV44255; QW44255; QX44255; QY44255; QZ44255; RA44255; RB44255; RC44255; RD44255; RE44255; RF44255; RG44255; RH44255; RI44255; RJ44255; RK44255; RL44255; RM44255; RN44255; RO44255; RP44255; RQ44255; RR44255; RS44255; RT44255; RU44255; RV44255; RW44255; RX44255; RY44255; RZ44255; SA44255; SB44255; SC44255; SD44255; SE44255; SF44255; SG44255; SH44255; SI44255; SJ44255; SK44255; SL44255; SM44255; SN44255; SO44255; SP44255; SQ44255; SR44255; SS44255; ST44255; SU44255; SV44255; SW44255; SX44255; SY44255; SZ44255; TA44255; TB44255; TC44255; TD44255; TE44255; TF44255; TG44255; TH44255; TI44255; TJ44255; TK44255; TL44255; TM44255; TN44255; TO44255; TP44255; TQ44255; TR44255; TS44255; TT44255; TU44255; TV44255; TW44255; TX44255; TY44255; TZ44255; UA44255; UB44255; UC44255; UD44255; UE44255; UF44255; UG44255; UH44255; UI44255; UJ44255; UK44255; UL44255; UM44255; UN44255; UO44255; UP44255; UQ44255; UR44255; US44255; UT44255; UU44255; UV44255; UW44255; UX44255; UY44255; UZ44255; VA44255; VB44255; VC44255; VD44255; VE44255; VF44255; VG44255; VH44255; VI44255; VJ44255; VK44255; VL44255; VM44255; VN44255; VO44255; VP44255; VQ44255; VR44255; VS44255; VT44255; VU44255; VV44255; VW44255; VX44255; VY44255; VZ44255; WA44255; WB44255; WC44255; WD44255; WE44255; WF44255; WG44255; WH44255; WI44255; WJ44255; WK44255; WL44255; WM44255; WN44255; WO44255; WP44255; WQ44255; WR44255; WS44255; WT44255; WU44255; WV44255; WW44255; WX44255; WY44255; WZ44255; XA44255; XB44255; XC44255; XD44255; XE44255; XF44255; XG44255; XH44255; XI44255; XJ44255; XK44255; XL44255; XM44255; XN44255; XO44255; XP44255; XQ44255; XR44255; XS44255; XT44255; XU44255; XV44255; XW44255; XX44255; XY44255; XZ44255; YA44255; YB44255; YC44255; YD44255; YE44255; YF44255; YG44255; YH44255; YI44255; YJ44255; YK44255; YL44255; YM44255; YN44255; YO44255; YP44255; YQ44255; YR44255; YS44255; YT44255; YU44255; YV44255; YW44255; YX44255; YY44255; YZ44255; ZA44255; ZB44255; ZC44255; ZD44255; ZE44255; ZF44255; ZG44255; ZH44255; ZI44255; ZJ44255; ZK44255; ZL44255; ZM44255; ZN44255; ZO44255; ZP44255; ZQ44255; ZR44255; ZS44255; ZT44255; ZU44255; ZV44255; ZW44255; ZX44255; ZY44255; ZZ44255;

R;Taylor, M.E.; Conary, J.T.; Lennartz, M.R.; Stahl, P.D.; Drickamer, K.

J. Biol. Chem. 265, 12156-12162, 1990

A;Title: Primary structure of the mannose receptor contains multiple motifs resembling ,

A;Reference number: A36563; MUID:90324192; PMID:2373685

A;Accession: A36563

A;Molecule type: mRNA

A;Residues: 1-1456 <TAY>

A;Cross-references: GB:J05550; NID:gl88675; PIDN:AAA59868.1; PID:gl88676

A;Note: parts of this sequence, including the amino end of the mature protein, were con

R;Ezekowitz, R.A.B.; Sastry, K.; Bailly, P.; Warner, A.

J. Exp. Med. 172, 1785-1794, 1990

A;Title: Molecular characterization of the human macrophage mannose receptor: demonstra

A;Reference number: A60926; MUID:91079783; PMID:2258707

A;Accession: A60926

A;Status: nucleic acid sequence not shown

A;Molecule type: mRNA

A;Residues: 1-1333,'T',1335-1456 <EZE>

A;Cross-references: GB:X55635

A;Note: translation of the nucleotide sequence is incomplete

A;Note: in the authors' translation additional residues Pro-Glu-Ile are shown after 497

R;Kim, S.J.; Ruiz, N.; Bezouska, K.; Drickamer, K.

Genomics 14, 721-727, 1992

A;Title: Organization of the gene encoding the human macrophage mannose receptor (MRC1)

A;Reference number: A44255; MUID:93052405; PMID:1294118

A;Accession: A44255

A;Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tr

A;Molecule type: DNA

A;Residues: 155-233,'KSAL',238-283;346-428;492-569;631-714,716-719;783-820,'N',822-865;

A;Note: sequence extracted from NCB1 backbone (NCBIP:118415, NCBIP:118421, NCBIP:118428

C;Genetics:

A;Gene: GDB:MRC1

A;Cross-references: GDB:133759; OMIM:153618

A;Map position: 10p13-10p13

C;Superfamily: phospholipase A2 receptor; C-type lectin homology; fibronectin type II r

C;Keywords: duplication; lectin; tandem repeat; transmembrane protein

F;1-18/Domain: signal sequence #status predicted <SIG>

F;168-209/Domain: fibronectin type II repeat homology <2F1>

F;223-340/Domain: C-type lectin homology <LCH1>

F;362-486/Domain: C-type lectin homology <LCH2>

F;945-1079/Domain: C-type lectin homology <LCH3>

Query Match

Best Local Similarity 16.3%; Score 182; DB 1; Length 1456;

Matches 52; Conservativity 37; Mismatches 61; Indels 54; Gaps 8;

QY 21 YKVIYFHDTSRRLLNFEEAKEACRRDGGQLVSIESTEQLIEKFIENLLPSDGFWIGLR 80

Db 807 YKDYYFSKEKETMDNARAFCKRNFGLVSIQSESEKKFLWKYV-NRNDQAQSAFYFICLL 865
QY 81 RREEKQSNSTACQDLYAWTDGSIQFRNWWYVDEPSCGS--EVCVVMYHQPAPAGIGGPY 138
Db 866 ISLDKK-----FAWMDGSKVDYVSWATGEPNFANEDENCVTMY-----SNSGF----- 908
QY 139 MFQWDDRCNMKNFICKYSDK----PAVPSREAEGEETELTTPVLPEETQE----- 187
Db 909 ---WNDINCGYPNAFICQRHNSINATTVM-----TPSPVSGCKEGWNPYSN 954
QY 188 -----EDAKKTFKESREAAAL 202
Db 955 KCFKIFGFMEERKNWQEARACI 978
RESULT 2
A48925
mannose receptor precursor, macrophage - mouse
C;Species: Mus musculus (house mouse)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A48925; S21320; PC2245
R;Harris, N.; Super, M.; Rits, M.; Chang, G.; Ezekowitz, R.A.
Blood 80, 2363-2373, 1992
A;Title: Characterization of the murine macrophage mannose receptor: demonstration that
A;Reference number: A48925; MUID:93043353; PMID:1421407
A;Accession: A48925
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-1455 <HAR>
A;Experimental source: peritoneal macrophage
A;Note: sequence extracted from NCBI backbone (NCBIP:118733)
R;Harris, N.; Super, M.; Rits, M.; Chang, G.; Ezekowitz, R.B.
submitted to the EMBL Data Library, April 1992
A;Description: Characterization of the murine macrophage mannose receptor: Demonstration
on.
A;Reference number: S21320
A;Accession: S21320
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-302,'W',303-1117,'E',1119-1455 <HA2>
A;Cross-references: EMBL:Z11974; NID:G52997; PIDN:CAA78028.1; PID:G52998
R;Harris, N.; Peters, L.L.; Eicher, E.M.; Rits, M.; Raspberry, D.; Eichbaum, Q.G.; Super
Biochem. Biophys. Res. Commun. 198, 682-692, 1994
A;Title: The exon-intron structure and chromosomal localization of the mouse macrophage
A;Reference number: PC2245; MUID:94128116; PMID:8297379
A;Accession: PC2245
A;Molecule type: mRNA
A;Residues: 35-105 <HA3>
C;Genetics:
A;Gene: Mrc1
A;Map position: 2
C;Superfamily: phospholipase A2 receptor; C-type lectin homology; fibronectin type II re
C;Keywords: membrane protein; receptor
F;168-209/Domain: fibronectin type II repeat homology <2F9>
F;361-485/Domain: C-type lectin homology <LCH1>
F;943-1077/Domain: C-type lectin homology <LCH2>
Query Match 16.0%; Score 178.5; DB 1; Length 1455;
Best Local Similarity 25.9%; Pred. No. 1.8e-07;
Matches 53; Conservative 33; Mismatches 66; Indels 53; Gaps 8;
QY 21 YKVIYFHDTSRRLNFEAKEACRRDGGQLVSIQSESEDEQKLIKFIENLLPSDGFWIGLR 80
Db 806 YKDYYFSKEKETMDNARRFCKNFGDLATIKSESEKKFLWKYI-NKNGGQSPYFIFGML 864
QY 81 RREEKQSNSTACQDLYAWTDGSIQFRNWWYVDEPSCGS--EVCVVMYHQPAPAGIGGPY 138
Db 865 ISMDKK-----FIWMDGSKVDVFAWATGEPNFANDDENCVTMY-----TNSGF----- 907
QY 139 MFQWDDRCNMKNFICK----YSDEKPAVPSREAEGEETELTTPVLPEETQE----- 187
Db 908 ---WNDINCGYPNPFICQRHNSINATAMP-----TTPTPGCKEGWHLKYNK 953

QY 188 -----EDAKKTFKESREAAALNL 204
Db 954 CPKIFGFANEKKSWQDARQACKGL 978
RESULT 3
S52781
neurocan - mouse
C;Species: Mus musculus (house mouse)
C;Date: 19-May-1995 #sequence_revision 21-Jul-1995 #text_change 04-Feb-2000
C;Accession: S52781
R;Rauch, U.; Forsberg, N.; Kulbe, G.; Arnold-Ammer, I.; Faessler, R.
submitted to the EMBL Data Library, February 1995
A;Description: Amino acid sequence of mouse neurocan and their different
A;Reference number: S52781
A;Accession: S52781
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1268 <RAU>
A;Cross-references: EMBL:X84727; NID:G758629; PIDN:CAA59216.1; PID:G758630
C;Superfamily: aggrecan; C-type lectin homology; complement factor H repeat
F;176-253/Domain: link protein repeat homology <LNK1>
F;274-355/Domain: link protein repeat homology <LNK2>
F;964-995/Domain: EGF homology <EGF>
F;1040-1160/Domain: C-type lectin homology <LCH>
F;1167-1223/Domain: complement factor H repeat homology <FHD>
Query Match 15.9%; Score 177; DB 2; Length 1268;
Best Local Similarity 31.5%; Pred. No. 2.1e-07;
Matches 45; Conservative 17; Mismatches 49; Indels 32; Gaps 6;
QY 17 QRPCYKVIYFHDTSRRLNFEAKEACRRDGGQLVSIQSESEDEQKLIKFIENLLPSDGFW 76
Db 1048 QGHCYR--YF---AHRRAWEDAERDCRRRAGHLTSVHSPEEHKFINSF-----GHNSW 1096
QY 77 IGLRRREKQSNSTACQDLYAWTDGSIQFRNWWYVDEPS---CGSEVCVVMYHQPAPAG 133
Db 1097 IGLNDRTVRD-----FQWTDNTGLQYENWREKQPDNFFAGGEDCVVMVAHESG--- 1145
QY 134 IGGPYMFQWDDRCNMKNFICK 156
Db 1146 -----RWNDVPCYNLPLPVCK 1161
RESULT 4
T14274
versican precursor, splice form V2 - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 05-May-2000
C;Accession: T14274
R;Schmalfeldt, M.; Dours-Zimmermann, M.T.; Winterhalter, K.H.; Zimmermann, D.R.
J. Biol. Chem. 273, 15758-15764, 1998
A;Title: Versican V2 is a major extracellular matrix component of the mature bovine bra
A;Reference number: Z17954; MUID:98288320; PMID:9624174
A;Accession: T14274
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1643 <SCH>
A;Cross-references: EMBL:AF060458; NID:G32533303; PID:G32533304; PIDN:AAC24360.1
A;Experimental source: brain
C;Keywords: glycoprotein
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-1643/Product: versican, splice form V2 #status predicted <MAT>
F;57,331,352,817,965,1017,1333,1616,1626/Binding site: carbohydrate (Asn) (covalent) #
Query Match 15.7%; Score 174.5; DB 2; Length 1643;
Best Local Similarity 25.5%; Pred. No. 4.8e-07;
Matches 59; Conservative 26; Mismatches 71; Indels 75; Gaps 11;
QY 17 QRPCYKVIYFHDTSRRLNFEAKEACRRDGGQLVSIQSESEDEQKLIKFIENLLPSDGF- 75
Db 1424 QGQCYK--YF---AHRRTWDAARECKLQGAHLTSILSHEEQMFVNRV-----GHQYQ 1471

RESULT 8
A60979
versican precursor - human
N:Alternate names: chondroitin sulfate proteoglycan 2; chondroitin sulfate proteoglycan
N:Contains: glial hyaluronate-binding protein
C:Species: Homo sapiens (man)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 19-Jan-2001
C:Accession: S06014; S43921; A60979; A30358; A29348; A45131; I54179
R:Zimmermann, D.R.; Ruoslahti, E.
EMBO J. 8, 2975-2981, 1989
A:Title: Multiple domains of the large fibroblast proteoglycan, versican.
A:Reference number: S06014; MUID:90059882; PMID:2583089
A:Accession: S06014
A:Molecule type: mRNA
A:Residues: 1-2409 <ZIM>
A:Cross-references: GB:X15998; NID:g37662; PIDN:CAA34128.1; PID:g37663
R:Yao, L.Y.; Moody, C.; Schoenherr, E.; Wight, T.N.; Sandell, L.J.
Matrix Biol. 14, 213-225, 1994
A:Title: Identification of the proteoglycan versican in aorta and smooth muscle cells by
A:Reference number: S43921; MUID:95005762; PMID:7921538
A:Accession: S43921
A:Molecule type: mRNA
A:Residues: 208-440;1094-1385;1910-2246 <YAO>
R:Bignami, A.; Lane, W.S.; Andrews, D.; Dahl, D.
Brain Res. Bull. 22, 67-70, 1989
A:Title: Structural similarity of hyaluronate binding proteins in brain and cartilage.
A:Reference number: A60979; MUID:89229983; PMID:2469524
A:Accession: A60979
A:Molecule type: protein
A:Residues: 171-210;289-303 <BIG>
R:Perides, G.; Lane, W.S.; Andrews, D.; Dahl, D.; Bignami, A.
J. Biol. Chem. 264, 5981-5987, 1989
A:Title: Isolation and partial characterization of a glial hyaluronate-binding protein.
A:Reference number: A30358; MUID:89174663; PMID:2466833
A:Accession: A30358
A:Molecule type: protein
A:Residues: 24-50;80-87,'D',89-119;128-155;167-218;229-259,'IR';261-268;277-283,'G',285-
R:Krusius, T.; Gehlsen, K.R.; Ruoslahti, E.
J. Biol. Chem. 262, 13120-13125, 1987
A:Title: A fibroblast chondroitin sulfate proteoglycan core protein contains lectin-like
A:Reference number: A29348; MUID:88007514; PMID:2820964
A:Accession: A29348
A:Molecule type: mRNA
A:Residues: 1725,'V',1727-2409 <KRU>
A:Cross-references: GB:J02814
R:Perides, G.; Rahemtulla, F.; Lane, W.S.; Asher, R.A.; Bignami, A.
J. Biol. Chem. 267, 23883-23887, 1992
A:Title: Isolation of a large aggregating proteoglycan from human brain.
A:Reference number: A45131; MUID:93054750; PMID:1429726
A:Contents: brain
A:Accession: A45131
A:Molecule type: protein
A:Residues: 21-22,'X',24-37 <PE2>
A:Experimental source: brain
A:Note: sequence extracted from NCBI backbone (NCBIP:118884)
R:Iozzo, R.V.; Naso, M.F.; Cannizzaro, L.A.; Wasmuth, J.J.; McPherson, J.D.
Genomics 14, 845-851, 1992
A:Title: Mapping of the versican proteoglycan gene (CSPG2) to the long arm of human chro
A:Reference number: I54179; MUID:93122792; PMID:1478664
A:Accession: I54179
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 251-347 <RES>
A:Cross-references: GB:S52488; NID:g263313; PIDN:AAB24878.1; PID:g263314
C:Genetics:
A:Gene: GDB:CSPG2
A:Cross-references: GDB:127873; OMIM:118661
A:Map position: 5q12-5q14
C:Superfamily: versican; C-type lectin homology; complement factor H repeat homology; EG
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-2409/Product: proteoglycan 24K core protein #status predicted <MAT>
F;167-244/Domain: link protein repeat homology <LNK1>

F;265-346/Domain: link protein repeat homology <LNK2>
F;539-1654/Domain: chondroitin sulfate attachment #status predicted <GAG>
F;2106-2137/Domain: EGF homology <EG1>
F;2144-2175/Domain: EGF homology <EG2>
F;2182-2302/Domain: C-type lectin homology <LCH>
F;2309-2365/Domain: complement factor H repeat homology <FHD>

Query Match 15.6%; Score 174; DB 1; Length 2409;
Best Local Similarity 28.5%; Pred. No. 8.3e-07;
Matches 47; Conservative 23; Mismatches 55; Indels 40; Gaps 8;

QY 17 QRPCYKVIYFHDTSRRLLNFEAEKACRRDGGQLVSI ESEDEQKLIKFIENLLPSD GDF- 75
| | | | | : : : : : | | | | | : : : : : | : : : : :
Db 2190 QGQCYK--YF---AHRRTWDAARECR LQGAHLTSILSHEEQVFN RV-----GHDYQ 2237
| | | | | : : : : : | | | | | : : : : : | : : : : :

QY 76 WIGLRRREEKQSNSTACQDLYAWTDGSI SQFRN WYVDEP-----SCGSEVCVVMYHQPSAP 131
| | | | | : : : : : | | | | | : : : : : | : : : : :
Db 2238 WIGL-----NDKMFEDFRWTDGST LQYENWRPNQPD SFFSAGEDCVV IWHENG-- 2287
| | | | | : : : : : | | | | | : : : : : | : : : : :

QY 132 AGIGGYPYMFQWDDRCNMKNFICKYS----DEKPAVPSREAE GE 172
| | | | | : : : : : | | | | | : : : : : | : : : : :
Db 2288 -----QWNDVPCNYHLTYTCKGTGTVACGQPPVVENAKTF GK 2323
| | | | | : : : : : | | | | | : : : : : | : : : : :

RESULT 9
A47171
chondroitin sulfate proteoglycan PG-M core protein - chicken
C:Species: Gallus gallus (Chicken)
C:Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 21-Jul-2000
C:Accession: A47171
R:Shinomura, T.; Nishida, Y.; Ito, K.; Kimata, K.
J. Biol. Chem. 268, 14461-14469, 1993
A:Title: cDNA cloning of PG-M, a large chondroitin sulfate proteoglycan expressed durin
A:Reference number: A47171; MUID:93300846; PMID:8314802
A:Accession: A47171
A:Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-3562 <SHI>
A:Cross-references: GB:D13542; NID:g391643; PIDN:BAA02742.1; PID:g391644
A:Experimental source: stage 22-23 developing limb buds
A:Note: sequence extracted from NCBI backbone (NCBIN:134456, NCBIP:134457)
C:Superfamily: chicken chondroitin sulfate proteoglycan PG-M core protein; C-type lectin
F;166-243/Domain: link protein repeat homology <LNK1>
F;264-345/Domain: link protein repeat homology <LNK2>
F;3258-3289/Domain: EGF homology <EGF1>
F;3296-3327/Domain: EGF homology <EGF>
F;3334-3454/Domain: C-type lectin homology <LCH>
F;3461-3517/Domain: complement factor H repeat homology <FHD>

Query Match 15.3%; Score 171; DB 2; Length 3562;
Best Local Similarity 28.5%; Pred. No. 2.4e-06;
Matches 47; Conservative 23; Mismatches 55; Indels 40; Gaps 8;

QY 17 QRPCYKVIYFHDTSRRLLNFEAEKACRRDGGQLVSI ESEDEQKLIKFIENLLPSD GDF- 75
| | | | | : : : : : | | | | | : : : : : | : : : : :
Db 3342 QGQCYK--YF---AHRRTWDAARECR LQGAHLTSILSHEEQVFN RI-----GHDYQ 3389
| | | | | : : : : : | | | | | : : : : : | : : : : :

QY 76 WIGLRRREEKQSNSTACQDLYAWTDGSI SQFRN WYVDEP-----SCGSEVCVVMYHQPSAP 131
| | | | | : : : : : | | | | | : : : : : | : : : : :
Db 3390 WIGL-----NDKMFEDFRWTDGST LQYENWRPNQPD SFFSAGEDCVV IWHENG-- 3439
| | | | | : : : : : | | | | | : : : : : | : : : : :

QY 132 AGIGGYPYMFQWDDRCNMKNFICKYS----DEKPAVPSREAE GE 172
| | | | | : : : : : | | | | | : : : : : | : : : : :
Db 3440 -----QWNDVPCNYHLTYTCKGTGTVACGQPPVVENAKTF GK 3475
| | | | | : : : : : | | | | | : : : : : | : : : : :

RESULT 10
T42710
mannose receptor, macrophage - mouse
N:Alternate names: lambda lectin; phospholipase A2 receptor
C:Species: Mus musculus (house mouse)
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jun-2000

C;Accession: T42710
R;Wu, K.; Yuan, J.; Lasky, L.A.
J. Biol. Chem. 271, 21323-21330, 1996
A;Title: Characterization of a novel member of the macrophage mannose receptor type C le
A;Reference number: Z22235; MUID:96355501; PMID:8702911
A;Accession: T42710
A;Status: preliminary; translated from GB/EMBL/DBBJ
A;Molecule type: mRNA
A;Residues: 1-1479 <WML>
A;Cross-references: EMBL:U56734; NID:G1336073; PID:G1336074; PIDN:AAC52729.1
C;Superfamily: phospholipase A2 receptor; C-type lectin homology; fibronectin type II re
C;Keywords: membrane protein; receptor
F;186-227/Domain: fibronectin type II repeat homology <2FR>

Query Match 15.3%; Score 170.5; DB 2; Length 1479;
Best Local Similarity 31.4%; Pred. No. 9.6e-07;
Matches 48; Conservative 22; Mismatches 48; Indels 35; Gaps 7;

QY 7 SQQVCRGGTQPCYKVIYFHDTSRRLNFEAEACRRDGGQLVSIIESEDEQKLIIEKFI 66
Db 384 SQQPF-----QGHCVRL-----QAERKSQWQESKRACLRGGDLLSIHSMAELEFITKQIK 433

QY 67 NLLPSDGFWIGLRRREKQSNSTACQDLYAWTDGSIQFRNWWYVDEPS---CGSEVCVV 123
Db 434 QEVE---ELWIGL-----NDLKLQMFWSGSLVSFTWHWPFEPNFRDLSLEDCVT 482

QY 124 MYHQPSAPAGIGGPFQWQNDRCNMKNFICK 156
Db 483 IW----GPEG-----RWNDSPCNQSLPSICK 504

RESULT 11
A39808
proteoglycan core protein, cartilage - bovine (fragments)
N;Alternate names: aggrecan; aggregating cartilage proteoglycan
C;Species: Bos primigenius taurus (cattle)
C;Date: 20-Mar-1992 #sequence revision 23-Mar-1995 #text change 13-Aug-1999
C;Accession: A34234; A27752; A39808; A27751; E29164; B27751; D27751; E27751; F27
R;Antonsson, P.; Heinegard, D.; Oldberg, A.
J. Biol. Chem. 264, 16170-16173, 1989
A;Title: The keratan sulfate-enriched region of bovine cartilage proteoglycan consists o
A;Reference number: A34234; MUID:89380219; PMID:2528543
A;Accession: A34234
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 128-621 <ANT>
A;Cross-references: GB:J05028
R;Oldberg, A.; Antonsson, P.; Heinegard, D.
Biochem. J. 243, 255-259, 1987
A;Title: The partial amino acid sequence of bovine cartilage proteoglycan, deduced from a
A;Reference number: A27752; MUID:87270630; PMID:3111460
A;Accession: A27752
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 622-1340 <OLD>
R;Sandy, J.D.; Boynton, R.E.; Flannery, C.R.
J. Biol. Chem. 266, 8198-8205, 1991
A;Title: Analysis of the catabolism of aggrecan in cartilage explants by quantitation of
A;Reference number: A39808; MUID:91217051; PMID:2022637
A;Accession: A39808
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-28;59-82;131-137,'QSET',142-149;196-207;226-249;1137-1143;1252-1267;1274-1
R;Perin, J.P.; Bonnet, F.; Jolles, P.
FEBS Lett. 206, 73-77, 1986
A;Title: Structural relationship between link proteins and proteoglycan monomers.
A;Reference number: A27751; MUID:87005253; PMID:3530809
A;Accession: A27751
A;Molecule type: protein
A;Residues: 29-58;74-130;174-175,'A',177-204;208-225 <PER>
R;Perin, J.P.; Bonnet, F.; Jolles, J.; Jolles, P.
FEBS Lett. 176, 37-42, 1984
A;Title: Sequence data concerning the protein core of the cartilage proteoglycan monomer

A;Reference number: A91327; MUID:85027710; PMID:6489519
A;Accession: E29164
A;Molecule type: protein
A;Residues: 1230-1249 <PE2>
C;Superfamily: aggrecan; C-type lectin homology; complement factor H repeat homology; EC
C;Keywords: glycoprotein
F;8-28/Domain: link protein repeat homology (fragment) <LNK1>
F;29-58/Domain: link protein repeat homology (fragment) <LNK2>
F;80-146/Domain: link protein repeat homology (fragments) <LNK3>
F;167-248/Domain: link protein repeat homology <LNK4>
F;1130-1250/Domain: C-type lectin homology <LCH>
F;1257-1313/Domain: complement factor H repeat homology <FHD>

Query Match 14.2%; Score 158.5; DB 2; Length 1340;
Best Local Similarity 26.8%; Pred. No. 9.9e-06;
Matches 48; Conservative 29; Mismatches 59; Indels 43; Gaps 10;

QY 9 QPVCRRGG---TQPCYKVIYFHDTSRRLNFEAEACRRDGGQLVSIIESEDEQKLIIEKFI 65
Db 1127 QKLCEEGWTKFQGHCYR--HFPD---RATWVDAESQCRKQSHLSSIVTPEEQ----EFV 1177

QY 66 ENLLPSDGF-WIGLRRREKQSNSTACQDLYAWTDGSIQFRNWWYVDEP----SCGSEV 120
Db 1178 NN---NAQDYQWIGL-----NDKTIEGDFRWSHGSLQFENWRPNQPDNFFATGEDC 1226

QY 121 CVVMYHQPSAPAGIGGPFQWQNDRCNMKNFICKYS----DEKPAVPSREAEGETE 175
Db 1227 VVMIWHEKG-----EWNDVPCNYQLPFTCKKGTACGEPVVEHARIFGQKKD 1274

RESULT 12
T42630
aggrecan - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000
C;Accession: T42630
R;Hering, T.M.; Kollar, J.; Huynh, T.D.
submitted to the EMBL Data Library, September 1996
A;Description: Complete coding sequence of bovine aggrecan: comparative structural analy
A;Reference number: Z22182
A;Accession: T42630
A;Status: preliminary; translated from GB/EMBL/DBBJ
A;Molecule type: mRNA
A;Residues: 1-2327 <HER>
A;Cross-references: EMBL:U76615; NID:G1730259; PID:G1730260; PIDN:AAB38524.1
A;Experimental source: articular chondrocytes
C;Superfamily: aggrecan; C-type lectin homology; complement factor H repeat homology; EC
C;Keywords: cartilage; chondroitin sulfate proteoglycan; extracellular matrix; glycoprot

Query Match 14.2%; Score 158.5; DB 2; Length 2327;
Best Local Similarity 26.8%; Pred. No. 1.9e-05;
Matches 48; Conservative 29; Mismatches 59; Indels 43; Gaps 10;

QY 9 QPVCRRGG---TQPCYKVIYFHDTSRRLNFEAEACRRDGGQLVSIIESEDEQKLIIEKFI 65
Db 2114 QKLCEEGWTKFQGHCYR--HFPD---RATWVDAESQCRKQSHLSSIVTPEEQ----EFV 2164

QY 66 ENLLPSDGF-WIGLRRREKQSNSTACQDLYAWTDGSIQFRNWWYVDEP----SCGSEV 120
Db 2165 NN---NAQDYQWIGL-----NDKTIEGDFRWSHGSLQFENWRPNQPDNFFATGEDC 2213

QY 121 CVVMYHQPSAPAGIGGPFQWQNDRCNMKNFICKYS----DEKPAVPSREAEGETE 175
Db 2214 VVMIWHEKG-----EWNDVPCNYQLPFTCKKGTACGEPVVEHARIFGQKKD 2261

RESULT 13
A39086
aggrecan precursor, cartilage long splice form [validated] - human
N;Alternate names: chondroitin sulfate proteoglycan 1; large aggregating proteoglycan; P
N;Contains: aggrecan cartilage short splice form
C;Species: Homo sapiens (man)
C;Date: 10-Sep-1999 #sequence_revision 01-Dec-2000 #text_change 08-Dec-2000

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 22, 2003, 16:07:10 ; Search time 19.6848 Seconds
(without alignments)
1954.412 Million cell updates/sec

Title: US-09-887-855-5
Perfect score: 1115
Sequence: 1 ATGRLLSGQPVCRGGTQRPC.....EEDAKKTFKESREAAALNLAY 206

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 696363 seqs, 186758610 residues

Total number of hits satisfying chosen parameters: 696363

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1115	100.0	206	9	US-09-887-855-5
2	1115	100.0	374	9	US-09-887-855-2
3	1115	100.0	374	12	US-10-094-749-2090
4	1115	100.0	374	15	US-10-149-819-15
5	1111	99.6	374	12	US-10-094-749-2142
6	1101	98.7	382	10	US-09-909-320-137
7	1101	98.7	382	10	US-09-909-088B-137
8	1101	98.7	382	10	US-09-905-291A-137
9	1101	98.7	382	10	US-09-902-853-137
10	1101	98.7	382	10	US-09-907-824-137
11	1101	98.7	382	10	US-09-907-841-137
12	1101	98.7	382	11	US-09-904-011-137
13	1101	98.7	382	11	US-09-906-742-137
14	1101	98.7	382	11	US-09-906-838-137
15	1101	98.7	382	11	US-09-907-613-137

16	1101	98.7	382	11	US-09-907-942-137	Sequence 137, App
17	1101	98.7	382	11	US-09-904-859-137	Sequence 137, App
18	1101	98.7	382	11	US-09-909-204-137	Sequence 137, App
19	1101	98.7	382	11	US-09-904-820-137	Sequence 137, App
20	1101	98.7	382	11	US-09-904-786-137	Sequence 137, App
21	1101	98.7	382	11	US-09-906-646-137	Sequence 137, App
22	1101	98.7	382	11	US-09-906-700-137	Sequence 137, App
23	1101	98.7	382	11	US-09-903-786-137	Sequence 137, App
24	1101	98.7	382	11	US-09-902-903-137	Sequence 137, App
25	1101	98.7	382	11	US-09-903-749A-137	Sequence 137, App
26	1101	98.7	382	11	US-09-904-119-137	Sequence 137, App
27	1101	98.7	382	11	US-09-904-956-137	Sequence 137, App
28	1101	98.7	382	11	US-09-902-736-137	Sequence 137, App
29	1101	98.7	382	11	US-09-907-794-137	Sequence 137, App
30	1101	98.7	382	11	US-09-903-943-137	Sequence 137, App
31	1101	98.7	382	11	US-09-904-462-137	Sequence 137, App
32	1101	98.7	382	11	US-09-907-925-137	Sequence 137, App
33	1101	98.7	382	11	US-09-902-692-137	Sequence 137, App
34	1101	98.7	382	11	US-09-903-520-137	Sequence 137, App
35	1101	98.7	382	11	US-09-905-056-137	Sequence 137, App
36	1101	98.7	382	11	US-09-909-064-137	Sequence 137, App
37	1101	98.7	382	11	US-09-904-553-137	Sequence 137, App
38	1101	98.7	382	11	US-09-905-381-137	Sequence 137, App
39	1101	98.7	382	11	US-09-905-088-137	Sequence 137, App
40	1101	98.7	382	11	US-09-907-575-137	Sequence 137, App
41	1101	98.7	382	11	US-09-905-075-137	Sequence 137, App
42	1101	98.7	382	11	US-09-902-759-137	Sequence 137, App
43	1101	98.7	382	11	US-09-902-634-137	Sequence 137, App
44	1101	98.7	382	11	US-09-902-713-137	Sequence 137, App
45	1101	98.7	382	11	US-09-907-979-137	Sequence 137, App

ALIGNMENTS

RESULT 1
US-09-887-855-5
; Sequence 5, Application US/09887855
; Patent No. US20020058310A1
; GENERAL INFORMATION:
; APPLICANT: Immunex Corporation
; APPLICANT: Anderson, Dirk M
; TITLE OF INVENTION: LECTIN SS3939 DNA AND POLYPEPTIDES
; FILE REFERENCE: 2883-US
; CURRENT APPLICATION NUMBER: US/09/887,855
; CURRENT FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 206
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-887-855-5

Query Match 100.0%; Score 1115; DB 9; Length 206;
Best Local Similarity 100.0%; Pred. No. 7.7e-105;
Matches 206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI	60	Sequence 5, Appli
Db	1	ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI	60	Sequence 2, Appli
QY	61	IEKFIENLLPSDGDWIGLRRRREKQSNSTACQDLYAWTDGSIQFRNMYVDEPSCGSEV	120	Sequence 2090, Ap
Db	61	IEKFIENLLPSDGDWIGLRRRREKQSNSTACQDLYAWTDGSIQFRNMYVDEPSCGSEV	120	Sequence 15, Appl
QY	121	CVVMYHOPSAPAGIGGYPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGETELTPV	180	Sequence 2142, Ap
Db	121	CVVMYHOPSAPAGIGGYPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGETELTPV	180	Sequence 137, App
QY	181	LPEETQEDAKKTFKESREAAALNLAY	206	Sequence 137, App
Db	181	LPEETQEDAKKTFKESREAAALNLAY	206	Sequence 137, App


```
RESULT 2
US-09-887-855-2
; Sequence 2, Application US/09887855
; Patent No. US20020058310A1
; GENERAL INFORMATION:
; APPLICANT: Immunex Corporation
; APPLICANT: Anderson, Dirk M
; TITLE OF INVENTION: LECTIN SS3939 DNA AND POLYPEPTIDES
; FILE REFERENCE: 2883-US
; CURRENT APPLICATION NUMBER: US/09/887,855
; CURRENT FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-887-855-2

Query Match      100.0%; Score 1115; DB 9; Length 374;
Best Local Similarity 100.0%; Pred. No. 1.7e-104;
Matches 206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI ESEDEQKL 60
   |||||
Db 22 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI ESEDEQKL 81

QY 61 IEKFIENLLPSDGDGFWIGLRRRREEKQSNSTACQDLYAWTDGSI S QFRNWWYVDEPSCGSEV 120
   |||||
Db 82 IEKFIENLLPSDGDGFWIGLRRRREEKQSNSTACQDLYAWTDGSI S QFRNWWYVDEPSCGSEV 141

QY 121 CVVMYHQPAPAGIGGPPYMFQWNDRCNMKNFICKYSDEKPAVPSREAEGETELTTPV 180
   |||||
Db 142 CVVMYHQPAPAGIGGPPYMFQWNDRCNMKNFICKYSDEKPAVPSREAEGETELTTPV 201

QY 181 LPEETQEDAKKTFKESREAAALNLAY 206
   |||||
Db 202 LPEETQEDAKKTFKESREAAALNLAY 227

RESULT 3
US-10-094-749-2090
; Sequence 2090, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
```

```
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2090
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-094-749-2090

Query Match      100.0%; Score 1115; DB 12; Length 374;
Best Local Similarity 100.0%; Pred. No. 1.7e-104;
Matches 206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI ESEDEQKL 60
   |||||
Db 22 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI ESEDEQKL 81

QY 61 IEKFIENLLPSDGDGFWIGLRRRREEKQSNSTACQDLYAWTDGSI S QFRNWWYVDEPSCGSEV 120
   |||||
Db 82 IEKFIENLLPSDGDGFWIGLRRRREEKQSNSTACQDLYAWTDGSI S QFRNWWYVDEPSCGSEV 141

QY 121 CVVMYHQPAPAGIGGPPYMFQWNDRCNMKNFICKYSDEKPAVPSREAEGETELTTPV 180
   |||||
Db 142 CVVMYHQPAPAGIGGPPYMFQWNDRCNMKNFICKYSDEKPAVPSREAEGETELTTPV 201

QY 181 LPEETQEDAKKTFKESREAAALNLAY 206
   |||||
Db 202 LPEETQEDAKKTFKESREAAALNLAY 227

RESULT 4
US-10-149-819-15
; Sequence 15, Application US/10149819
; Publication No. US20030044913A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: YUE, Henry
; APPLICANT: AZIMZAI, Valda
; APPLICANT: TANG, Y. Tom
; APPLICANT: PATTERSON, Chandra
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: SHAH, Purvi
; APPLICANT: LAL, Preeti
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: BURFORD, Neil
; TITLE OF INVENTION: EXTRACELLULAR MATRIX AND CELL ADHESION MOLECULES
; FILE REFERENCE: PF-0760 PCT
; CURRENT APPLICATION NUMBER: US/10/149,819
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/172,852; 60/172,354
; PRIOR FILING DATE: 1999-12-10; 1999-12-16
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PERL program
; SEQ ID NO 15
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030044913A1 3143411CD1
US-10-149-819-15

Query Match      100.0%; Score 1115; DB 15; Length 374;
Best Local Similarity 100.0%; Pred. No. 1.7e-104;
Matches 206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI ESEDEQKL 60
   |||||
Db 22 ATGRLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI ESEDEQKL 81

QY 61 IEKFIENLLPSDGDGFWIGLRRRREEKQSNSTACQDLYAWTDGSI S QFRNWWYVDEPSCGSEV 120
   |||||
Db 82 IEKFIENLLPSDGDGFWIGLRRRREEKQSNSTACQDLYAWTDGSI S QFRNWWYVDEPSCGSEV 141
```

QY 121 CVVMYHQPSAPAGIGPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGETELTPV 180
Db 142 CVVMYHQPSAPAGIGPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGETELTPV 201
QY 181 LPEETQEDAKKTFKESREAAALNLAY 206
Db 202 LPEETQEDAKKTFKESREAAALNLAY 227

RESULT 5

US-10-094-749-2142
; Sequence 2142, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHIKO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2142
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-094-749-2142

Query Match 99.6%; Score 1111; DB 12; Length 374;
Best Local Similarity 99.5%; Pred. No. 4.4e-104;
Matches 205; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGRLLSGQPVCRCGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI ESEDEQKL 60
Db 22 ATGRLLSGQPVCRCGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI ESEDEQKL 81
QY 61 IEKFIENLLPSDGFWIGLRRREEKQSNSTACQDLYAWTDGSI SQFRNMYVDEPSCGSEV 120
Db 82 IEKFIENLLPSDGFWIGLRRREEKQSNSTACQDLYAWTDGSI SQFRNMYVDEPSCGSEV 141
QY 121 CVVMYHQPSAPAGIGPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGETELTPV 180
Db 142 CVVMYHQPSAPAGIGPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGETELTPV 201
QY 181 LPEETQEDAKKTFKESREAAALNLAY 206
Db 202 LPEETQEDTKKTFKESREAAALNLAY 227

RESULT 6

US-09-909-320-137
; Sequence 137, Application US/09909320

; Patent No. US20020132240A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/909,320
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-909-320-137

Query Match 98.7%; Score 1101; DB 10; Length 382;
Best Local Similarity 96.3%; Pred. No. 4.6e-103;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

Qy 1 ATGRLLS-----GQVCRGGTQPCYKVIYFHDTSRRLNFEAEKACRRDGGQLVSI 52
Db 22 ATGRLLSASDLDLRGGQPVCRGGTQPCYKVIYFHDTSRRLNFEAEKACRRDGGQLVSI 81

Qy 53 ESEDEQKLIKFIENLLPSDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNWWYVD 112
Db 82 ESEDEQKLIKFIENLLPSDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNWWYVD 141

Qy 113 EPSCGSEVCVMYHQPSAPAGIGGYPMFQWNDRCNMKNFNICKYSDEKPAVPSREAEGE 172
Db 142 EPSCGSEVCVMYHQPSAPAGIGGYPMFQWNDRCNMKNFNICKYSDEKPAVPSREAEGE 201

Qy 173 ETELTTPVLPEETQEEADAKKTFKESREAAALNLAY 206
Db 202 ETELTTPVLPEETQEEADAKKTFKESREAAALNLAY 235

RESULT 7
US-09-909-088B-137
; Sequence 137, Application US/09909088B
; Patent No. US20020146709A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/909,088B
; CURRENT FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089

; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-909-088B-137

Query Match 98.7%; Score 1101; DB 10; Length 382;
Best Local Similarity 96.3%; Pred. No. 4.6e-103;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

Qy 1 ATGRLLS-----GQVCRGGTQPCYKVIYFHDTSRRLNFEAEKACRRDGGQLVSI 52
Db 22 ATGRLLSASDLDLRGGQPVCRGGTQPCYKVIYFHDTSRRLNFEAEKACRRDGGQLVSI 81

Qy 53 ESEDEQKLIKFIENLLPSDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNWWYVD 112
Db 82 ESEDEQKLIKFIENLLPSDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNWWYVD 141

Qy 113 EPSCGSEVCVMYHQPSAPAGIGGYPMFQWNDRCNMKNFNICKYSDEKPAVPSREAEGE 172
Db 142 EPSCGSEVCVMYHQPSAPAGIGGYPMFQWNDRCNMKNFNICKYSDEKPAVPSREAEGE 201

Qy 173 ETELTTPVLPEETQEEADAKKTFKESREAAALNLAY 206
Db 202 ETELTTPVLPEETQEEADAKKTFKESREAAALNLAY 235

RESULT 8
US-09-905-291A-137
; Sequence 137, Application US/09905291A
; Patent No. US20020160374A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel

```

; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/905,291A
; CURRENT FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-905-291A-137

Query Match      98.7%; Score 1101; DB 10; Length 382;
Best Local Similarity 96.3%; Pred. No. 4.6e-103;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY      1  ATGRLLS-----GQVCRGGTQPCYKVIYFHDTSRRRLNFEEAKEACRRDGGQLVSI 52
Db      1  |||||
Db      22  ATGRLLSASDLDLRGGQPVCRGGTQPCYKVIYFHDTSRRRLNFEEAKEACRRDGGQLVSI 81
QY      53  ESEDEQKLIKFIENLLPSDGDGFWIGLRRRREEKQSNSTACQDLYAWTDGSIQFRNWWYVD 112
Db      53  |||||
Db      82  ESEDEQKLIKFIENLLPSDGDGFWIGLRRRREEKQSNSTACQDLYAWTDGSIQFRNWWYVD 141
QY      113  EPSCGSEVCVMYHQPAPAGIGGPFYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGE 172
Db      113  |||||
Db      142  EPSCGSEVCVMYHQPAPAGIGGPFYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGE 201
QY      173  ETELTTPVLPEETQEEADAKKTFKESREAAALNLAY 206
Db      173  |||||
Db      202  ETELTTPVLPEETQEEADAKKTFKESREAAALNLAY 235

RESULT 9
US-09-902-853-137
; Sequence 137, Application US/09902853
; Publication No. US20020192659A1
```

```

; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/902,853
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US/09/665,350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-902-853-137
```

Query Match

98.7%; Score 1101; DB 10; Length 382;

Best Local Similarity 96.3%; Pred. No. 4.6e-103;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 ATGRLLS-----GQVCRGGTQRPCYKVIYFHDTSRRLNFEFEEAKEACRRDGGQLVSI 52
|||||
Db 22 ATGRLLSASDLRLGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEFEEAKEACRRDGGQLVSI 81

QY 53 ESEDEQKLIKFIENLLPSDGDGFWIGLRRRREEKQSNSTACQDLYAWTDGSGISQFRNWWYVD 112
|||||
Db 82 ESEDEQKLIKFIENLLPSDGDGFWIGLRRRREEKQSNSTACQDLYAWTDGSGISQFRNWWYVD 141

QY 113 EPSCGSEVCVMYHQPSAPAGIGGYPYMFQWNDRCNMKNNFICKYSDEKPAVPSREAAGE 172
|||||
Db 142 EPSCGSEVCVMYHQPSAPAGIGGYPYMFQWNDRCNMKNNFICKYSDEKPAVPSREAAGE 201

QY 173 ETELTTPVLPEETQEEADAKKTFKESREAAALNLAY 206
|||||
Db 202 ETELTTPVLPEETQEEADAKKTFKESREAAALNLAY 235

RESULT 10
US-09-907-824-137
; Sequence 137, Application US/09907824
; Publication No. US20020197671A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/907,824
; CURRENT FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: 09/665,350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15

; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO: 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-907-824-137

Query Match 98.7%; Score 1101; DB 10; Length 382;
Best Local Similarity 96.3%; Pred. No. 4.6e-103;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 ATGRLLS-----GQVCRGGTQRPCYKVIYFHDTSRRLNFEFEEAKEACRRDGGQLVSI 52
|||||
Db 22 ATGRLLSASDLRLGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEFEEAKEACRRDGGQLVSI 81

QY 53 ESEDEQKLIKFIENLLPSDGDGFWIGLRRRREEKQSNSTACQDLYAWTDGSGISQFRNWWYVD 112
|||||
Db 82 ESEDEQKLIKFIENLLPSDGDGFWIGLRRRREEKQSNSTACQDLYAWTDGSGISQFRNWWYVD 141

QY 113 EPSCGSEVCVMYHQPSAPAGIGGYPYMFQWNDRCNMKNNFICKYSDEKPAVPSREAAGE 172
|||||
Db 142 EPSCGSEVCVMYHQPSAPAGIGGYPYMFQWNDRCNMKNNFICKYSDEKPAVPSREAAGE 201

QY 173 ETELTTPVLPEETQEEADAKKTFKESREAAALNLAY 206
|||||
Db 202 ETELTTPVLPEETQEEADAKKTFKESREAAALNLAY 235

RESULT 11
US-09-907-841-137
; Sequence 137, Application US/09907841
; Publication No. US20020198366A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.

```

; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/907,841
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-907-841-137

Query Match          98.7%; Score 1101; DB 10; Length 382;
Best Local Similarity 96.3%; Pred. No. 4.6e-103;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

Qy 1 ATGRLLS-----GQVCRGGTQPCYKVIYFHDTSRRLLNFEEAKEACRRDGGQLVSI 52
Db 22 ATGRLLSASDLRLGGQPCVCRGGTQPCYKVIYFHDTSRRLLNFEEAKEACRRDGGQLVSI 81

Qy 53 ESEDEQKLIKFIENLLPSDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNWWYVD 112
Db 82 ESEDEQKLIKFIENLLPSDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNWWYVD 141

Qy 113 EPSCGSEVCVMYHQSPAPAGIGGPFYQWDDRCNMKNFNICKYSDEKPAVPSREAEGE 172
Db 142 EPSCGSEVCVMYHQSPAPAGIGGPFYQWDDRCNMKNFNICKYSDEKPAVPSREAEGE 201

Qy 173 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 206
Db 202 ETELTTPVLPEETQEEDAKKTFKESREAALNLAY 235

RESULT 12
US-09-904-011-137
; Sequence 137, Application US/09904011
; Publication No. US20030003530A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
```

```

; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/904,011
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 09/665,350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-904-011-137
```

```

Query Match          98.7%; Score 1101; DB 11; Length 382;
Best Local Similarity 96.3%; Pred. No. 4.6e-103;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

Qy 1 ATGRLLS-----GQVCRGGTQPCYKVIYFHDTSRRLLNFEEAKEACRRDGGQLVSI 52
Db 22 ATGRLLSASDLRLGGQPCVCRGGTQPCYKVIYFHDTSRRLLNFEEAKEACRRDGGQLVSI 81

Qy 53 ESEDEQKLIKFIENLLPSDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNWWYVD 112
Db 82 ESEDEQKLIKFIENLLPSDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNWWYVD 141
```

QY	113	EPSCGSEVCVVMYHQPSPAPAGIGGYPYMFQWDDRCNMKNFICKYSDCKPAVPSPRAEAGE	172
Db	142	EPSCGSEVCVVMYHQPSPAPAGIGGYPYMFQWDDRCNMKNFICKYSDCKPAVPSPRAEAGE	201
QY	173	ETELTTPVLPEETQEEEDAKKTFKESREAAINLAY	206
Db	202	ETELTTPVLPEETQEEEDAKKTFKESREAAINLAY	235

RESULT 13

US-09-906-742-137
; Sequence 137, Application US/09906742
; Publication No. US20030023054A1
; GENERAL INFORMATION:

```

; PRIOR APPLICATION NUMBER: PCT/US99/300995
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/309111
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/309999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/002199
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-906-742-137

```

Query Match	98.7%;	Score 1101;	DB 11;	Length 382;
Best Local Similarity	96.3%;	Pred. No. 4.6e-103;		
Matches 206;	Conservative	0;	Mismatches	0;
			Indels	8

QY	1	ATGRLLS-----GQVCRGGTQRPCYKVIYFHDTSRRLNFEAAKEACRRDGGQLVSI	52
Db	22	ATGRLLSASDLRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEAAKEACRRDGGQLVSI	81
QY	53	ESEDEQKLIKFIEIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNWWYVD	112
Db	82	ESEDEQKLIKFIEIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNWWYVD	141
QY	113	EPSCGSEVCVMYHQPSAPAGIGGPPYMFQWDDRCNMKNFICKYSDEKPAVPSREAAGE	172
Db	142	EPSCGSEVCVMYHQPSAPAGIGGPPYMFQWDDRCNMKNFICKYSDEKPAVPSREAAGE	201
QY	173	ETELTTPVLPEETQEEADAKKTFKESREAAALNLAY	206
Db	202	ETELTTPVLPEETQEEADAKKTFKESREAAALNLAY	235

RESULT 14

US-09-906-838-137
; Sequence 137, Application US/09906838
; Publication No. US20030027143A1

; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-906-838-137

Query Match 98.7%; Score 1101; DB 11; Length 382;
Best Local Similarity 96.3%; Pred. No. 4.6e-103;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 ATGRLLS-----GQPVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
Db 22 ATGRLLSASDLRLGGQPVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

QY 53 ESEDEQKLIETIENLLPSDGFWIGLRRRREKQSNSTACQDLYAWTDGSIQFRNWWVD 112
Db 82 ESEDEQKLIETIENLLPSDGFWIGLRRRREKQSNSTACQDLYAWTDGSIQFRNWWVD 141

QY 113 EPSCGSEVCVMYHOPSAPAGIGGYPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGE 172
Db 142 EPSCGSEVCVMYHOPSAPAGIGGYPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGE 201

QY 173 ETELTTPVLPEETQEDAKKTFKESREAAALNLAY 206
Db 202 ETELTTPVLPEETQEDAKKTFKESREAAALNLAY 235

RESULT 15
US-09-907-613-137
; Sequence 137, Application US/09907613
; Publication No. US20030027145A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone

; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/907,613
; PRIOR FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-907-613-137

Query Match 98.7%; Score 1101; DB 11; Length 382;
Best Local Similarity 96.3%; Pred. No. 4.6e-103;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 ATGRLLS-----GQPVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
Db 22 ATGRLLSASDLRLGGQPVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

Qy	53	ESEDEQKLI	EKFIENLL	PSDGF	WIGLR	RRREK	QSNST	ACODL	YAWTD	GSISQ	FRN	WYVD	112
Db	82	ESEDEQKLI	EKFIENLL	PSDGF	WIGLR	RRREK	QSNST	ACODL	YAWTD	GSISQ	FRN	WYVD	141
Qy	113	EPSCGSE	VCV	MYH	QPS	APAGI	GGP	YMFQ	WDD	RCN	MKN	FN	172
Db	142	EPSCGSE	VCV	MYH	QPS	APAGI	GGP	YMFQ	WDD	RCN	MKN	FN	201
Qy	173	ETELT	TPVL	PEET	QEE	DAK	KTF	KES	REAA	LN	LAY		206
Db	202	ETELT	TPVL	PEET	QEE	DAK	KTF	KES	REAA	LN	LAY		235

Search completed: December 22, 2003, 16:16:49
Job time : 20.6848 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 22, 2003, 16:05:09 ; Search time 6.46605 Seconds
(without alignments)
1476.391 Million cell updates/sec

Title: US-09-887-855-2_COPY_25_227

Perfect score: 1100

Sequence: 1 RLLSGPVCRCGGTQPCYKV.....EEDAKTFKESREALNLAY 203

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	574	52.2	273	1	CHOD_MOUSE
2	561.5	51.0	273	1	CHOD_HUMAN
3	182	16.5	1456	1	MANR_HUMAN
4	177	16.1	1268	1	PGCN_MOUSE
5	176	16.0	1321	1	PGCN_HUMAN
6	174.5	15.9	3381	1	PGCV_BOVIN
7	174	15.8	1257	1	PGCN_RAT
8	174	15.8	2738	1	PGCV_RAT
9	174	15.8	3358	1	PGCV_MOUSE
10	174	15.8	3396	1	PGCV_HUMAN
11	171	15.5	3562	1	PGCV_CHICK
12	164	14.9	643	1	CD93_RAT
13	158.5	14.4	2364	1	PGCA_BOVIN
14	158.5	14.4	2415	1	PGCA_HUMAN
15	155.5	14.1	2333	1	PGCA_CANFA
16	154.5	14.0	612	1	LEM2_MOUSE
17	153.5	14.0	644	1	CD93_MOUSE
18	153.5	14.0	2124	1	PGCA_RAT
19	152	13.8	912	1	PGCB_BOVIN
20	151.5	13.8	652	1	CD93_HUMAN
21	151	13.7	197	1	CLF1_HUMAN
22	149	13.5	321	1	FCE2_HUMAN
23	147.5	13.4	2132	1	PGCA_MOUSE
24	147	13.4	883	1	PGCB_MOUSE
25	146	13.3	2109	1	PGCA_CHICK
26	145	13.2	158	1	LECG_TRIST
27	145	13.2	883	1	PGCB_RAT
28	144.5	13.1	173	1	LEC2_MEGRO
29	144.5	13.1	372	1	LEM1_RAT
30	143.5	13.0	372	1	LEM1_MOUSE
31	141.5	12.9	331	1	FCE2_MOUSE
32	141.5	12.9	549	1	LEM2_RAT
33	140.5	12.8	162	1	LEC3_MEGRO
					Q9cxm0 mus musculus
					Q9h9p2 homo sapien
					P22897 homo sapien
					P55066 mus musculus
					O14594 homo sapien
					P81282 bos taurus
					P55067 rattus norv
					Q9erb4 rattus norv
					Q62059 mus musculus
					P13611 homo sapien
					Q90953 gallus gall
					Q9et61 rattus norv
					P13608 bos taurus
					P16112 homo sapien
					Q28343 canis famil
					Q00690 mus musculus
					O89103 mus musculus
					P07897 rattus norv
					Q28062 bos taurus
					Q9npy3 homo sapien
					O75596 homo sapien
					P06734 homo sapien
					Q61282 mus musculus
					Q61361 mus musculus
					P07898 gallus gall
					Q9ygp1 trimeresuru
					P55068 rattus norv
					P17346 megabalanus
					P30836 rattus norv
					P18337 mus musculus
					P20693 mus musculus
					P98105 rattus norv
					P07439 megabalanus

ALIGNMENTS

RESULT 1

CHOD_MOUSE

ID	CHOD_MOUSE	STANDARD;	PRT;	273 AA.
AC	Q9CXM0; Q8VI31;			
DT	28-FEB-2003 (Rel. 41, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Chondrolectin precursor (Transmembrane protein MT75).			
GN	CHODL.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=C57BL/6J;			
RA	Weng L., Smits P., Hubner R., Wouters J., Merregaert J.;			
RT	"Mt75, a low expressed c-type lectin gene involving in			
RT	chondrogenesis."			
RL	Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=C57BL/6J; TISSUE=Embryonic head;			
RX	MEDLINE=21085660; PubMed=11217851;			
RA	Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,			
RA	Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,			
RA	Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,			
RA	Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,			
RA	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,			
RA	Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,			
RA	Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,			
RA	Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,			
RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,			
RA	Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,			
RA	Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,			
RA	Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,			
RA	Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombarts P.,			
RA	Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,			
RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,			
RA	Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,			
RA	Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,			
RA	Hayashizaki Y.;			
RT	"Functional annotation of a full-length mouse cDNA collection.";			
RL	Nature 409:685-690(2001).			
CC	-1- SUBCELLULAR LOCATION: Type I membrane protein (Potential).			
CC	-1- SIMILARITY: Contains 1 C-type lectin family domain.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; AF311699; AAL50354.1; -.			

P98131 bos taurus
P07714 homo sapien
P05047 sarcophaga
P23806 trimeresuru
P43025 mus musculus
Q95198 macaca mula
Q28768 papio hamad
P23132 bos taurus
P43137 mus musculus
P83300 anser anser
Q95235 pongo pygma
P98107 bos taurus

34 140 12.7 370 1 LEM1_BOVIN
35 139 12.6 248 1 PSPA_HUMAN
36 139 12.6 283 1 LECA_SARPE
37 138.5 12.6 152 1 IXA_TRIFL
38 138.5 12.6 202 1 TETN_MOUSE
39 137.5 12.5 372 1 LEM1_PAPHA
40 137.5 12.5 372 1 LEM1_BOVIN
41 136 12.4 175 1 LITH_BOVIN
42 135 12.3 165 1 LIT1_MOUSE
43 134.5 12.2 132 1 ACAL_ANSEAN
44 134.5 12.2 372 1 LEM1_PONPY
45 134.5 12.2 485 1 LEM2_BOVIN

DR	SMART; SM00059; FN2; 1.	
DR	SMART; SM00458; RICIN; 1.	
DR	PROSITE; PS00615; C_TYPE_LLECTIN_1; 6.	
DR	PROSITE; PS00041; C_TYPE_LLECTIN_2; 8.	
DR	PROSITE; PS00023; FIBRONECTIN_2; 1.	
DR	PROSITE; PS00231; RICIN_B_LLECTIN; 1.	
KW	Receptor; Signal; Calcium-binding; Transmembrane; Repeat;	
KW	Glycoprotein; Endocytosis; Macrophage; Lectin; Antigen; 3D-structure.	
FT	SIGNAL 1 18	POTENTIAL.
FT	CHAIN 19 1456	MACROPHAGE MANNOSE RECEPTOR.
FT	DOMAIN 19 1383	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 1384 1411	POTENTIAL.
FT	DOMAIN 1412 1456	CYTOPLASMIC (POTENTIAL).
FT	DOMAIN 22 142	RICIN B-TYPE LECTIN.
FT	DOMAIN 157 212	FIBRONECTIN TYPE-II.
FT	DOMAIN 216 344	C-TYPE LECTIN 1 (LONG FORM).
FT	DOMAIN 360 490	C-TYPE LECTIN 2 (LONG FORM).
FT	DOMAIN 502 629	C-TYPE LECTIN 3 (LONG FORM).
FT	DOMAIN 644 781	C-TYPE LECTIN 4 (LONG FORM).
FT	DOMAIN 805 926	C-TYPE LECTIN 5 (LONG FORM).
FT	DOMAIN 943 1083	C-TYPE LECTIN 6 (LONG FORM).
FT	DOMAIN 1100 1216	C-TYPE LECTIN 7 (LONG FORM).
FT	DOMAIN 1228 1359	C-TYPE LECTIN 8 (LONG FORM).
FT	DISULFID 646 659	
FT	DISULFID 680 777	
FT	DISULFID 753 769	
FT	CARBOHYD 104 104	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 344 344	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 529 529	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 926 926	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 930 930	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 1160 1160	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 1205 1205	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 1311 1311	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	TURN 648 649	
FT	STRAND 651 652	
FT	TURN 654 655	
FT	STRAND 658 663	
FT	HELIX 667 669	
FT	STRAND 671 671	
FT	HELIX 673 683	
FT	TURN 684 684	
FT	STRAND 686 687	
FT	HELIX 693 705	
FT	TURN 706 707	
FT	TURN 709 710	
FT	STRAND 712 718	
FT	TURN 723 724	
FT	STRAND 727 727	
FT	TURN 729 730	
FT	STRAND 733 733	
FT	TURN 741 742	
FT	HELIX 746 748	
FT	STRAND 752 757	
FT	TURN 758 761	
FT	STRAND 764 768	
FT	TURN 769 770	
FT	STRAND 773 780	
FT	TURN 781 782	
SQ	SEQUENCE 1456 AA; 166011 MW; 264E5AF3C576A5E3 CRC64;	
Query Match		
Best Local Similarity 16.5%; Score 182; DB 1; Length 1456;		
Matches 52; Conservative 37; Mismatches 61; Indels 54; Gaps 8;		
QY	18 YKVIYFHTSRRLNFEAKEACRRDGGQLVSIIESEDEQKLEKFIENLLPSDGFWIGLR 77	
Db	807 YKDYQYFYSKEKETMDNARAFCKRNFGLVSIQSEKKFLWKYV-NRNDQAQSAFYIGLL 865	
QY	78 RREEKQSNSTACQDLYAWTDGSIISQFRNWWYVDEPSCGS--EVCVVMYHQPSAPAGIGGPY 135	
Db	866 ISLDKK-----FAWMDGSKVDYVSWATGEPNFANEDENCVTMY-----SNSGF---- 908	

QY	136 MFQWNDRCNMKNFNICKYSDEK-----PAVPSREAGEBETELTTPVLPEETQE----- 184	
Db	909 ---WNDINCGYPNAPICQRHNSSINATTVM-----TMPSPVSGCKEGWNPYSN 954	
QY	185 -----EDAKTKFKESREAAAL 199	
Db	955 KCPKIFGFMEERKNWQEARACI 978	
RESULT 4		
PGCN_MOUSE		
ID	PGCN_MOUSE	STANDARD; PRT; 1268 AA.
AC	P55056;	
DT	01-OCT-1996 (Rel. 34, Created)	
DT	01-OCT-1996 (Rel. 34, Last sequence update)	
DT	15-SEP-2003 (Rel. 42, Last annotation update)	
DE	Neurocan core protein precursor (Chondroitin sulfate proteoglycan 3).	
GN	CSPG3 OR NCAN.	
OS	Mus musculus (Mouse).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
OX	NCBI_TaxID=10090;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=BALB/c; TISSUE=Brain;	
RX	MEDLINE=96039250; PubMed=7490074;	
RA	Rauch U., Grimpe B., Kulbe G., Arnold-Ammer I., Beier D.,	
RA	Fassler R.;	
RT	"Structure and chromosomal localization of the mouse neurocan gene.";	
RL	Genomics 28:405-410(1995).	
CC	-!- FUNCTION: May modulate neuronal adhesion and neurite growth during	
CC	development by binding to neural cell adhesion molecules (NG-CAM	
CC	and N-CAM). Chondroitin sulfate proteoglycan; binds to hyaluronic	
CC	acid.	
CC	-!- TISSUE SPECIFICITY: BRAIN.	
CC	-!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.	
CC	-!- SIMILARITY: Contains 2 EGF-like domains.	
CC	-!- SIMILARITY: Contains 2 link domains.	
CC	-!- SIMILARITY: Contains 1 C-type lectin family domain.	
CC	-!- SIMILARITY: Contains 1 Sushi (SCR) domain.	
CC	-!- SIMILARITY: BELONGS TO THE AGGRECAN/VERSICAN PROTEOGLYCAN FAMILY.	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation	
CC	the European Bioinformatics Institute. There are no restrictions on its	
CC	use by non-profit institutions as long as its content is in no way	
CC	modified and this statement is not removed. Usage by and for commercial	
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/	
CC	or send an email to license@isb-sib.ch).	
CC	-----	
DR	EMBL; X84727; CAA59216.1; --	
DR	PIR; S52781; S52781.	
DR	HSSP; P00740; 1EDM.	
DR	MGD; MGI:104694; Cspg3.	
DR	InterPro; IPR002353; AntifreezeII.	
DR	InterPro; IPR000152; Asx hydroxyl.	
DR	InterPro; IPR000742; EGF_2.	
DR	InterPro; IPR001881; EGF_Ca.	
DR	InterPro; IPR006209; EGF_like.	
DR	InterPro; IPR007110; Ig-Like.	
DR	InterPro; IPR003599; Ig.	
DR	InterPro; IPR003006; Ig_MHC.	
DR	InterPro; IPR001304; Lectin_C.	
DR	InterPro; IPR000538; Link.	
DR	InterPro; IPR000436; Sushi_SCR_CCP.	
DR	Pfam; PF00008; EGF; 2.	
DR	Pfam; PF00047; ig; 1.	
DR	Pfam; PF00059; lectin_c; 1.	
DR	Pfam; PF00084; sushi; 1.	
DR	Pfam; PF00193; Xlink; 2.	
DR	PRINTS; PR01265; LINKMODULE.	
DR	PRINTS; PR00356; ANTIFREEZEII.	
DR	ProDom; PD000918; Link; 2.	

DR	SMART; SM00032; CCP; 1.	
DR	SMART; SM00034; CLECT; 1.	
DR	SMART; SM00179; EGF_CA; 1.	
DR	SMART; SM00409; IG; 1.	
DR	SMART; SM00445; LINK; 2.	
DR	PROSITE; PS00010; ASX_HYDROXYL; 1.	
DR	PROSITE; PS00615; C_TYPE_LLECTIN_1; 1.	
DR	PROSITE; PS50041; C_TYPE_LLECTIN_2; 1.	
DR	PROSITE; PS00022; EGF_1; 3.	
DR	PROSITE; PS01186; EGF_2; 1.	
DR	PROSITE; PS01187; EGF_CA; 1.	
DR	PROSITE; PS50835; IG_LIKE; 1.	
DR	PROSITE; PS01241; LINK; 2.	
KW	Glycoprotein; Hyaluronic acid; Proteoglycan; Immunoglobulin domain;	
KW	EGF-like domain; Calcium; Repeat; Lectin; Sushi; Signal.	
FT	SIGNAL 1 22 POTENTIAL.	
FT	CHAIN 23 1268 NEUROCAN CORE PROTEIN.	
FT	DOMAIN 37 157 IG-LIKE V-TYPE.	
FT	DOMAIN 158 253 LINK 1.	
FT	DOMAIN 259 355 LINK 2.	
FT	DOMAIN 960 996 EGF-LIKE 1.	
FT	DOMAIN 998 1034 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).	
FT	DOMAIN 1036 1165 C-TYPE LECTIN.	
FT	DOMAIN 1166 1224 SUSHI.	
FT	DISULFID 58 139 BY SIMILARITY.	
FT	DISULFID 181 252 BY SIMILARITY.	
FT	DISULFID 205 226 BY SIMILARITY.	
FT	DISULFID 279 354 BY SIMILARITY.	
FT	DISULFID 303 324 BY SIMILARITY.	
FT	DISULFID 964 975 BY SIMILARITY.	
FT	DISULFID 969 984 BY SIMILARITY.	
FT	DISULFID 986 995 BY SIMILARITY.	
FT	DISULFID 1040 1051 BY SIMILARITY.	
FT	DISULFID 1068 1160 BY SIMILARITY.	
FT	DISULFID 1136 1152 BY SIMILARITY.	
FT	DISULFID 1167 1210 BY SIMILARITY.	
FT	DISULFID 1196 1223 BY SIMILARITY.	
FT	CARBOHYD 121 121 N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD 742 742 N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD 978 978 N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD 1175 1175 N-LINKED (GLCNAC. . .) (POTENTIAL).	
SQ	SEQUENCE 1268 AA; 137200 MW; 30148E202A2FAEC CRC64;	
Query Match 16.1%; Score 177; DB 1; Length 1268;		
Best Local Similarity 31.5%; Pred. No. 7.9e-08;		
Matches 45; Conservative 17; Mismatches 49; Indels 32; Gaps 6;		
Qy	14 QRPCYVIYFHDTSRRLLNFEEAKEACRRDGGQLVSISEDEKLIKFIENLLPSDGF 73	
Db	1048 QGHCVY--YF---AHRRRAWDAERDCRRRAGHLTSVHSPEEHKFINSF-----GHENSW 1096	
Qy	74 IGLRREEKQSNSTACQDLYAWTDGSIQFRNWWYVDEPS---CGSEVCVMYHQPSAPAG 130	
Db	1097 IGLNDRTVRD-----FQWTDNTGLQYENWREKQPDNFFAGGDCVMVAHESG--- 1145	
Qy	131 IGGPYMFQWDDRCNMKNFICK 153	
Db	1146 -----RWNDVPCNYNLPYCK 1161	
RESULT 5		
PGCN_HUMAN		
ID	PGCN_HUMAN STANDARD; PRT; 1321 AA.	
AC	O14594; Q9UPK6;	
DT	28-FEB-2003 (Rel. 41, Created)	
DT	28-FEB-2003 (Rel. 41, Last sequence update)	
DT	15-SEP-2003 (Rel. 42, Last annotation update)	
DE	Neurocan core protein precursor (Chondroitin sulfate proteoglycan 3).	
GN	CSPG3 OR NCAN OR NEUR.	
OS	Homo sapiens (Human).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	

OX	NCBI_TaxID=9606;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RX	MEDLINE=99013874; PubMed=9795216;	
RA	Prange C.K., Pennacchio L.A., Lieuallen K., Fan W., Lennon G.G.;	
RT	"Characterization of the human neurocan gene, CSPG3.";	
RL	Gene 221:199-205(1998).	
RN	[2]	
RP	SEQUENCE OF 1-990 AND 1007-1321 FROM N.A.	
RA	Lamerdin J.E., McCready P.M., Skowronski E., Adamson A.W.,	
RA	Burkhardt-Schultz K., Gordon L., Kyle A., Ramirez M., Stilwagen S.,	
RA	Phan H., Velasco N., Do L., Regala W., Terry A., Garnes J.,	
RA	Danganan L., Poundstone P., Christensen M., Georgescu A., Avila J.,	
RA	Liu S., Attix C., Andreise T., Frankheim M., Amico-Keller G.,	
RA	Coefield J., Duarte S., Lucas S., Bruce R., Thomas P., Quan G.,	
RA	Kronmiller B., Arellano A., Montgomery M., Ow D., Nolan M., Trong S.,	
RA	Kobayashi A., Olsen A.S., Carrano A.V.;	
RT	"Sequence analysis of an ~1 Mb region containing the MEF2B gene in	
RT	19p12.";	
RL	Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.	
CC	-1- FUNCTION: May modulate neuronal adhesion and neurite growth during	
CC	development by binding to neural cell adhesion molecules (NG-CAM	
CC	and N-CAM). Chondroitin sulfate proteoglycan; binds to hyaluronic	
CC	acid.	
CC	-1- TISSUE SPECIFICITY: Brain.	
CC	-1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.	
CC	-1- SIMILARITY: Contains 2 EGF-like domains.	
CC	-1- SIMILARITY: Contains 2 link domains.	
CC	-1- SIMILARITY: Contains 1 C-type lectin family domain.	
CC	-1- SIMILARITY: Contains 1 Sushi (SCR) domain.	
CC	-1- SIMILARITY: BELONGS TO THE AGGRECAN/VERSICAN PROTEOGLYCAN FAMILY.	
CC	-----	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -	
CC	the European Bioinformatics Institute. There are no restrictions on its	
CC	use by non-profit institutions as long as its content is in no way	
CC	modified and this statement is not removed. Usage by and for commercial	
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/	
CC	or send an email to license@isb-sib.ch).	
CC	-----	
DR	EMBL; AF026547; AAC80576.1; -	
DR	EMBL; AC003110; AAB86655.1; -	
DR	EMBL; AC005254; AAC25581.1; -	
DR	HSSP; P00740; 1EDM.	
DR	Genew; HGNC:2465; CSPG3.	
DR	MIM; 600826; -	
DR	InterPro; IPR000152; Asx_hydroxyl.	
DR	InterPro; IPR000742; EGF_2.	
DR	InterPro; IPR001881; EGF_Ca.	
DR	InterPro; IPR006209; EGF_like.	
DR	InterPro; IPR007110; IG-Like.	
DR	InterPro; IPR003599; IG.	
DR	InterPro; IPR003006; IG_MHC.	
DR	InterPro; IPR001304; Lectin_C.	
DR	InterPro; IPR000538; Link.	
DR	InterPro; IPR000436; Sushi_SCR_CCP.	
DR	Pfam; PF00008; EGF; 2.	
DR	Pfam; PF00047; ig; 1.	
DR	Pfam; PF00059; lectin_c; 1.	
DR	Pfam; PF00084; sushi; 1.	
DR	Pfam; PF00193; xlink; 2.	
DR	PRINTS; PR01265; LINKMODULE.	
DR	ProDom; PD000918; Link; 2.	
DR	SMART; SM00032; CCP; 1.	
DR	SMART; SM00034; CLECT; 1.	
DR	SMART; SM00179; EGF_CA; 1.	
DR	SMART; SM00409; IG; 1.	
DR	SMART; SM00445; LINK; 2.	
DR	PROSITE; PS00010; ASX_HYDROXYL; 1.	
DR	PROSITE; PS00615; C_TYPE_LLECTIN_1; 1.	
DR	PROSITE; PS50041; C_TYPE_LLECTIN_2; 1.	
DR	PROSITE; PS00022; EGF_1; 3.	
DR	PROSITE; PS01186; EGF_2; 1.	

DR Pfam; PF00193; Xlink; 2.
DR PRINTS; PR01265; LINKMODULE.
DR PRINTS; PR00356; ANTIFREEZEII.
DR ProDom; PD000918; Link; 2.
DR SMART; SM00032; CCP; 1.
DR SMART; SM00034; CLECT; 1.
DR SMART; SM00179; EGF CA; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00445; LINK; 2.
DR PROSITE; PS00010; ASX HYDROXYL; 1.
DR PROSITE; PS00615; C TYPE LECTIN 1; 1.
DR PROSITE; PS50041; C TYPE LECTIN 2; 1.
DR PROSITE; PS00022; EGF 1; 2.
DR PROSITE; PS01186; EGF 2; 1.
DR PROSITE; PS01187; EGF CA; 1.
DR PROSITE; PS50835; IG LIKE; 1.
DR PROSITE; PS01241; LINK; 2.
KW Glycoprotein; Lectin; Extracellular matrix; Sushi;
KW Signal; Repeat; EGF-like domain; Calcium; Immunoglobulin domain;
KW Hyaluronic acid; Alternative splicing.
FT SIGNAL 1 20
FT CHAIN 21 3381
FT DOMAIN 21 147
FT DOMAIN 168 245
FT DOMAIN 266 347
FT DOMAIN 349 1336
FT DOMAIN 1337 3074
FT DOMAIN 3074 3110
FT DOMAIN 3112 3148
FT DOMAIN 3161 3275
FT DOMAIN 3280 3338
FT DISULFID 44 131
FT DISULFID 173 244
FT DISULFID 197 218
FT DISULFID 271 346
FT DISULFID 295 316
FT DISULFID 3078 3089
FT DISULFID 3083 3098
FT DISULFID 3100 3109
FT DISULFID 3116 3127
FT DISULFID 3121 3136
FT DISULFID 3138 3147
FT DISULFID 3154 3165
FT DISULFID 3182 3274
FT DISULFID 3250 3266
FT DISULFID 3281 3324
FT DISULFID 3310 3337
FT CARBOHYD 57 57
FT CARBOHYD 331 331
FT CARBOHYD 352 352
FT CARBOHYD 817 817
FT CARBOHYD 965 965
FT CARBOHYD 1017 1017
FT CARBOHYD 1333 1333
FT CARBOHYD 1393 1393
FT CARBOHYD 1437 1437
FT CARBOHYD 1463 1463
FT CARBOHYD 1653 1653
FT CARBOHYD 1974 1974
FT CARBOHYD 2045 2045
FT CARBOHYD 2074 2074
FT CARBOHYD 2103 2103
FT CARBOHYD 2263 2263
FT CARBOHYD 2290 2290
FT CARBOHYD 2356 2356
FT CARBOHYD 2623 2623
FT CARBOHYD 2641 2641
FT CARBOHYD 2919 2919
FT CARBOHYD 3052 3052
FT CARBOHYD 3354 3354
FT CARBOHYD 3364 3364
FT VARSPLIC 349 349
P -> R (in isoform V1 and isoform V3).

FT VARSPLIC 350 1336 /FTid=VSP_003078.
FT Missing (in isoform V1).
FT VARSPLIC 1337 3074 /FTid=VSP_003079.
FT Missing (in isoform V2).
FT VARSPLIC 350 3074 /FTid=VSP_003080.
FT Missing (in isoform V3).
FT CONFLICT 25 25 /FTid=VSP_003081.
FT CONFLICT 51 51 MISSING (IN REF. 2).
FT CONFLICT 89 89 MISSING (IN REF. 2).
FT CONFLICT 96 96 N -> D (IN REF. 2).
FT CONFLICT 346 346 Q -> D (IN REF. 2).
FT CONFLICT 3381 AA; 369984 MW; F09716FA7778D459 CRC64; C -> R (IN REF. 2).
SQ SEQUENCE 3381 AA; 369984 MW; F09716FA7778D459 CRC64;
Query Match 15.9%; Score 174.5; DB 1; Length 3381;
Best Local Similarity 25.5%; Pred. No. 4.4e-07;
Matches 59; Conservative 26; Mismatches 71; Indels 75; Gaps 11;
QY 14 QRPCYKVIYFHDTSRRLLNFEEAEACRRDGGQLVSISEDEQKLEKFIENLLPSDGD- 72
Db 3162 QGQCYK--YF---AHRRTWDAARECLQGAHLTSILSHEEQMFVNRV-----GHDYQ 3209
QY 73 WIGLRRREEKQSNSTACQDLYAWTDCGISQFRNMYVDEP---SCGSEVCVVMYHQPSAP 128
Db 3210 WIGL-----NDKMFHDFRWTGSLQYENWRPNQPDSPFFSTGDCVVIWHENG-- 3259
QY 129 AGIGGYPYMFQWDDRCNMKNFICKYS-----DEKPAVPSREAE----- 169
Db 3260 -----QWMDVPCNYHLTYTCKKGTGTVACGQPPVVENAKTFGKMKPRYEINSLIRYHC 3310
QY 170 -----ETELTT-----PVL-----PEETQEEDAKKTPKESREAAALN 200
Db 3311 KDGFIQRHLPTIRCLNGRWMAMPKITCLNPSAYQRTYSKKYFKNSSSAKDN 3361
RESULT 7
PGCN RAT
ID_PGCN RAT STANDARD; PRT; 1257 AA.
AC P55067;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Neurocan core protein precursor (Chondroitin sulfate proteoglycan 3)
DE (245 kDa early postnatal core glycoprotein) [Contains: 150 kDa adult
DE core glycoprotein].
GN CSPG3 OR NCAN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=Sprague-Dawley; TISSUE=Brain;
RX MEDLINE=92406907; PubMed=1326557;
RA Rauch U., Karthikeyan L., Maurel P., Margolis R.U., Margolis R.K.;
RT "Cloning and primary structure of neurocan, a developmentally
RT regulated, aggregating chondroitin sulfate proteoglycan of brain.";
RL J. Biol. Chem. 267:19536-19547(1992).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=94230574; PubMed=7513709;
RA Friedlander D.R., Milev P., Karthikeyan L., Margolis R.K.,
RA Margolis R.U., Grumet M.;
RT "The neuronal chondroitin sulfate proteoglycan neurocan binds to the
RT neural cell adhesion molecules Ng-CAM/L1/NILS and N-CAM, and inhibits
RT neuronal adhesion and neurite outgrowth.";
RL J. Cell Biol. 125:669-680(1994).
CC -!- FUNCTION: May modulate neuronal adhesion and neurite growth during
CC development by binding to neural cell adhesion molecules (NG-CAM
CC and N-CAM). Chondroitin sulfate proteoglycan; binds to hyaluronic
CC acid.
CC -!- TISSUE SPECIFICITY: EARLY POSTNATAL AND ADULT BRAIN; NOT EXPRESSED
CC IN KIDNEY, LUNG, LIVER AND MUSCLE.

CC -!- PTM: CONTAINS MOSTLY CHONDROITIN SULFATE, BUT ALSO N-LINKED AND
CC O-LINKED OLIGOSACCHARIDES (BY SIMILARITY).
CC -!- PTM: TWO ISOFORMS WERE FOUND THAT PROBABLY ARISE BY PROTEOLYTIC
CC DEGRADATION. THE LARGE ISOFORM IS PREDOMINANT IN EARLY POSTNATAL
CC BRAIN, THE SMALL ISOFORM IS FOUND IN ADULT BRAIN.
CC -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
CC -!- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
CC -!- SIMILARITY: CONTAINS 2 LINK DOMAINS.
CC -!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 SUSHI (SCR) DOMAIN.
CC -!- SIMILARITY: BELONGS TO THE AGGREGAN/VERSICAN PROTEOGLYCAN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M97161; AAC37679.1; -.
CC PIR; S28764; S28764.
CC HSSP; P00740; 1EDM.
CC InterPro; IPR002353; AntifreezeII.
CC InterPro; IPR000152; Asx hydroxyl.
CC InterPro; IPR000742; EGF_2.
CC InterPro; IPR001881; EGF_Ca.
CC InterPro; IPR006209; EGF_like.
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR003599; Ig.
CC InterPro; IPR003006; Ig_MHC.
CC InterPro; IPR001304; Lectin_C.
CC InterPro; IPR000538; Link.
CC InterPro; IPR000436; Sushi_SCR_CCP.
CC Pfam; PF00008; EGF; 2.
CC Pfam; PF00047; ig; 1.
CC Pfam; PF00059; lectin_c; 1.
CC Pfam; PF00084; sushi; 1.
CC Pfam; PF00193; Xlink; 2.
CC PRINTS; PR01265; LINKMODULE.
CC PRINTS; PR00356; ANTIFREEZEII.
CC ProDom; PD000918; Link; 2.
CC SMART; SM00032; CCP; 1.
CC SMART; SM00034; CLECT; 1.
CC SMART; SM00179; EGF_CA; 1.
CC SMART; SM00409; IG_1.
CC SMART; SM00445; LINK; 2.
CC PROSITE; PS00010; ASX_HYDROXYL; 1.
CC PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
CC PROSITE; PS00041; C_TYPE_LECTIN_2; 1.
CC PROSITE; PS00022; EGF_1; 3.
CC PROSITE; PS01186; EGF_2; 1.
CC PROSITE; PS01187; EGF_CA; 1.
CC PROSITE; PS00835; IG_LIKE; 1.
CC PROSITE; PS01241; LINK; 2.
KW Glycoprotein; Hyaluronic acid; Proteoglycan; Immunoglobulin domain;
KW EGF-like domain; Calcium; Repeat; Lectin; Sushi; Signal.
FT SIGNAL 1 22
FT CHAIN 23 1257 NEUROCAN CORE PROTEIN.
FT CHAIN 639 1257 150 kDa ADULT CORE GLYCOPROTEIN.
FT DOMAIN 37 157 IG-LIKE V-TYPE.
FT DOMAIN 158 253 LINK 1.
FT DOMAIN 259 355 LINK 2.
FT DOMAIN 949 985 EGF-LIKE 1.
FT DOMAIN 987 1023 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 1025 1154 C-TYPE LECTIN.
FT DOMAIN 1155 1213 SUSHI.
FT DISULFID 58 139 BY SIMILARITY.
FT DISULFID 181 252 BY SIMILARITY.
FT DISULFID 205 226 BY SIMILARITY.
FT DISULFID 279 354 BY SIMILARITY.
FT DISULFID 303 324 BY SIMILARITY.
FT DISULFID 953 964 BY SIMILARITY.

FT DISULFID 958 973 BY SIMILARITY.
FT DISULFID 975 984 BY SIMILARITY.
FT DISULFID 1029 1040 BY SIMILARITY.
FT DISULFID 1057 1149 BY SIMILARITY.
FT DISULFID 1125 1141 BY SIMILARITY.
FT DISULFID 1156 1199 BY SIMILARITY.
FT DISULFID 1185 1212 BY SIMILARITY.
FT CARBOHYD 121 121 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 339 339 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 737 737 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 944 944 O-LINKED (XYL. .) (CHONDROITIN SULFATE).
FT CARBOHYD 967 967 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1164 1164 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 1257 AA; 135544 MW; 9928333DCFA19EE1B CRC64;
Query Match 15.8%; Score 174; DB 1; Length 1257;
Best Local Similarity 30.8%; Pred. No. 1.5e-07;
Matches 44; Conservative 18; Mismatches 49; Indels 32; Gaps 6;
QY 14 QRPCYKVIYFHDTSRRLNFEAEKACRRDGGQGLVLSIESEDEQKLIKFIENLLPSDGF 73
Db 1037 QGHCYR--YF---AHRRAWEDAERDCRRRAGHLTSVHSPEHKFINSP-----CHENSW 1085
QY 74 IGLRRREKQSNSTACQDLYAWTDGSIQFRNMYVDEPS---CGSEVCVMYTHQPSAPAG 130
Db 1086 IGLNDRTVERD-----FQWTDNTGLQYENWREKQPDNFFAGGEDCVMVVAHENG--- 1134
QY 131 IGGPYMFQWNNDRRCNMKNFICK 153
Db 1135 -----RWNDVPCNYPYVCK 1150
RESULT 8
PGCV RAT
ID_PGCV RAT STANDARD; PRT; 2738 AA.
AC Q9ERB4; O08592; O88564; Q9R1K4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Versican core protein precursor (Large fibroblast proteoglycan)
DE (Chondroitin sulfate proteoglycan core protein 2) (PG-M) (Glial
DE hyaluronate-binding protein) (GHAP) (Fragments).
UN CP5G2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE OF 349-2738 FROM N.A. (ISOFORM V0), SEQUENCE FROM N.A.
RP (ISOFORM V3), AND SEQUENCE OF 2657-2738 FROM N.A. (ISOFORM VINT).
RC STRAIN=Wistar Kyoto;
RX MEDLINE=99327053; PubMed=10397680;
RA Lemire J.M., Braun K.R., Maurel P., Kaplan E.D., Schwartz S.M.,
RA Wight T.N.;
RT "Versican/PG-M isoforms in vascular smooth muscle cells.";
RL Arterioscler. Thromb. Vasc. Biol. 19:1630-1639(1999).
RN [2]
RP SEQUENCE OF 349-2738 FROM N.A. (ISOFORM V0).
RC STRAIN=Wistar Kyoto;
RX MEDLINE=98308094; PubMed=9642104;
RA Milev P., Maurel P., Chiba A., Mevissen M., Popp S., Yamaguchi Y.,
RA Margolis R.K., Margolis R.U.;
RT "Differential regulation of expression of hyaluronan-binding
RT proteoglycans in developing brain: aggrecan, versican, neurocan, and
RT brevican.";
RL Biochem. Biophys. Res. Commun. 247:207-212(1998).
RN [3]
RP SEQUENCE OF 2421-2463 FROM N.A. (ISOFORM V0).
RC TISSUE=Kidney;
RX MEDLINE=98094159; PubMed=9434070;
RA Pyke C., Kristensen P., Ostergaard P.B., Oturai P.S., Romer J.;
RT "Proteoglycan expression in the normal rat kidney.";
RL Nephron 77:461-470(1997).

RESULT 9
PGCV_MOUSE STANDARD; PRT; 3358 AA.
AC Q62059; Q62058; Q9CUU0;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Versican core protein precursor (Large fibroblast proteoglycan)
DE (Chondroitin sulfate proteoglycan core protein 2) (PG-M).
GN CSPG2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS V0; V1 AND V2).
RC STRAIN=C57BL/6, and Swiss Webster; TISSUE=Brain;
RX MEDLINE=95122551; PubMed=7822336;
RA Ito K., Shinomura T., Zako M., Ujita M., Kimata K.;
RT "Multiple forms of mouse PG-M, a large chondroitin sulfate
RT proteoglycan generated by alternative splicing."
RL J. Biol. Chem. 270:958-965(1995).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM V3).
RC STRAIN=C57BL/6;
RX MEDLINE=95181355; PubMed=7876137;
RA Zako M., Shinomura T., Ujita M., Ito K., Kimata K.;
RT "Expression of PG-M(V3), an alternatively spliced form of PG-M
RT without a chondroitin sulfate attachment in region in mouse and human
RT tissues."
RL J. Biol. Chem. 270:3914-3918(1995).
RN [3]
RP SEQUENCE OF 1-1692 FROM N.A. (ISOFORM V1).
RC STRAIN=C57BL/6J; TISSUE=Skin;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Bromstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sabaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wyszaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
RN [4]
RP INTERACTION WITH FBLN1.
RX PubMed=10400671;
RA Aspborg A., Adam S., Kostka G., Timpl R., Heinegaard D.;
RT "Fibulin-1 is a ligand for the C-type lectin domains of aggrecan and
RT versican."
RL J. Biol. Chem. 274:20444-20449(1999).
CC -!- FUNCTION: May play a role in intercellular signaling and in
CC connecting cells with the extracellular matrix. May take part in
CC the regulation of cell motility, growth and differentiation. Binds
CC hyaluronic acid.
CC -!- SUBUNIT: Interacts with FBLN1.
CC -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=4;
CC Comment=Additional isoforms seem to exist;
CC Name=V0;
CC IsoId=Q62059-1; Sequence=Displayed;

CC Name=V1;
CC IsoId=Q62059-2; Sequence=VSP_003087, VSP_003088;
CC Name=V2;
CC IsoId=Q62059-3; Sequence=VSP_003089;
CC Name=V3;
CC IsoId=Q62059-4; Sequence=VSP_003087, VSP_003090;
CC TISSUE SPECIFICITY: V2 is found only in brain.
CC -!- DEVELOPMENTAL STAGE: Disappears after the cartilage development.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -!- SIMILARITY: Contains 2 link domains.
CC -!- SIMILARITY: Contains 2 EGF-like domains.
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC -!- SIMILARITY: Contains 1 Sushi (SCR) domain.
CC -!- SIMILARITY: BELONGS TO THE AGGREGAN/VERSICAN PROTEOGLYCAN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D16263; BAA03796.1; -;
CC EMBL; D28599; -; NOT ANNOTATED_CDS.
CC EMBL; D32040; BAA06802.1; -;
CC EMBL; AK014525; BAB29411.1; -;
CC HSSP; P01132; 1EPG.
CC MGD; MGI:102889; Cspg2.
CC InterPro; IPR000152; Asx hydroxyl.
CC InterPro; IPR000742; EGF_2.
CC InterPro; IPR001881; EGF_Ca.
CC InterPro; IPR006209; EGF-like.
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR003599; Ig.
CC InterPro; IPR003006; Ig_MHC.
CC InterPro; IPR001304; Lectin_C.
CC InterPro; IPR000538; Link.
CC InterPro; IPR000436; Sushi_SCR_CCP.
CC Pfam; PF00008; EGF; 2.
CC Pfam; PF00047; Ig; 1.
CC Pfam; PF00059; lectin_c; 1.
CC Pfam; PF00084; sushi; 1.
CC Pfam; PF00193; Xlink; 2.
CC PRINTS; PR01265; LINKMODULE.
CC ProDom; PD000918; Link; 2.
CC SMART; SM00032; CCP; 1.
CC SMART; SM00034; CLECT; 1.
CC SMART; SM00179; EGF_CA; 1.
CC SMART; SM00409; IG; 1.
CC SMART; SM00445; LINK; 2.
CC PROSITE; PS00010; ASX HYDROXYL; 1.
CC PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
CC PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
CC PROSITE; PS00022; EGF_1_2.
CC PROSITE; PS01186; EGF_2; 1.
CC PROSITE; PS01187; EGF_CA; 1.
CC PROSITE; PS50835; IG_LIKE; 1.
CC PROSITE; PS01241; LINK; 2.
CC Glycoprotein; Proteoglycan; Lectin; Extracellular matrix; Sushi;
CC Signal; Repeat; EGF-like domain; Calcium; Immunoglobulin domain;
CC Hyaluronic acid; Alternative splicing.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 3358 VERSICAN CORE PROTEIN.
FT DOMAIN 21 146 IG-LIKE V-TYPE.
FT DOMAIN 167 244 LINK 1.
FT DOMAIN 265 346 LINK 2.
FT DOMAIN 348 1308 GAG-ALPHA
(GLUCOSAMINOGLYCAN ATTACHMENT DOMAIN).
FT DOMAIN 1309 3052 GAG-BETA.
FT DOMAIN 3052 3088 EGF-LIKE 1.
FT DOMAIN 3090 3126 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 3139 3253 C-TYPE LECTIN.


```
FT DOMAIN 3258 3316 SUSHI.
FT DISULFID 44 130 BY SIMILARITY.
FT DISULFID 172 243 BY SIMILARITY.
FT DISULFID 196 217 BY SIMILARITY.
FT DISULFID 270 333 BY SIMILARITY.
FT DISULFID 294 315 BY SIMILARITY.
FT DISULFID 3056 3067 BY SIMILARITY.
FT DISULFID 3061 3076 BY SIMILARITY.
FT DISULFID 3078 3087 BY SIMILARITY.
FT DISULFID 3094 3105 BY SIMILARITY.
FT DISULFID 3099 3114 BY SIMILARITY.
FT DISULFID 3116 3125 BY SIMILARITY.
FT DISULFID 3132 3143 BY SIMILARITY.
FT DISULFID 3160 3252 BY SIMILARITY.
FT DISULFID 3228 3244 BY SIMILARITY.
FT DISULFID 3259 3302 BY SIMILARITY.
FT DISULFID 3288 3315 BY SIMILARITY.
FT CARBOHYD 57 57 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 330 330 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 351 351 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 441 441 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 807 807 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 914 914 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 951 951 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1305 1305 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1372 1372 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1679 1679 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2054 2054 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2244 2244 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2362 2362 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2627 2627 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 3030 3030 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 3332 3332 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 3342 3342 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 348 348 P -> R (in isoform V1 and isoform V3).
FT VARSPLIC 349 1308 Missing (in isoform V1).
FT VARSPLIC 1309 3052 Missing (in isoform V2).
FT VARSPLIC 349 3052 Missing (in isoform V3).
FT CONFLICT 126 126 A -> G (IN REF. 3).
FT CONFLICT 348 348 MISSING (IN REF. 3).
FT CONFLICT 1658 1658 I -> T (IN REF. 3).
FT CONFLICT 1674 1680 TVWNSNS -> QFGIQTA (IN REF. 3).
SQ SEQUENCE 3358 AA; 366938 MW; 071B80026BC0762D CRC64;

Query Match 15.8%; Score 174; DB 1; Length 3358;
Best Local Similarity 28.5%; Pred. No. 4.9e-07;
Matches 47; Conservative 23; Mismatches 55; Indels 40; Gaps 8;

QY 14 QRPCYKVIYFHDTSRRLNPEEAKACRRDGGQGVLSIESEDEQKLEKFIENLLPSDGF- 72
Db 3140 QGCYK--YF---AHRRTWDAARECRLOGAHLTSILSHEEQMFVNRV-----GHDYQ 3187

QY 73 WIGLRREEKQSNSTACQDLYAWTDGSGISQFRNVYVDEP-----SCGSEVCVVMYHQPSAP 128
Db 3188 WIGL-----NDKMFEDHDFRWTGDSALQYENWRPNQPDSPFFSAGEDCVVIWHENG-- 3237

QY 129 AGIGPYMFQWNDRCNMKNFICKYS----DEKPAVPSREAGE 169
Db 3238 -----QWNVPCNYHLTYTCKKGTGTVACGQPPVVENAKTFGK 3273

RESULT 10
PGCV HUMAN
ID PGCV HUMAN STANDARD; PRT; 3396 AA.
AC P13611; P20754; Q13010; Q13189; Q15123; Q9UNW5;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Versican core protein precursor (Large fibroblast proteoglycan)
```

```
DE (Chondroitin sulfate proteoglycan core protein 2) (PG-M) (Glial
DE hyaluronate-binding protein) (GHAP).
GN CSPG2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM V0).
RX MEDLINE=95105188; PubMed=7528742;
RA Naso M.F., Zimmermann D.R., Iozzo R.V.;
RT "Characterization of the complete genomic structure of the human
RL versican gene and functional analysis of its promoter.";
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM V1).
RX TISSUE=Placenta;
RA Zimmermann D.R., Ruoslahti E.;
RT "Multiple domains of the large fibroblast proteoglycan, versican.";
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM V2).
RX TISSUE=Glial tumor;
RA Dours-Zimmermann M.T., Zimmermann D.R.;
RT "A novel glycosaminoglycan attachment domain identified in two
RL alternative splice variants of human versican.";
RN [4]
RP SEQUENCE OF 2711-3396 FROM N.A.
RX TISSUE=Lung fibroblast;
RA Iozzo R.V., Naso M.F., Cannizzaro L.A., Wasmuth J.J.,
RT "Mapping of the versican proteoglycan gene (CSPG2) to the long arm of
RL human chromosome 5 (5q12-5q14).";
RN [6]
RP SEQUENCE FROM N.A. (ISOFORM V3).
RX TISSUE=Brain;
RA Zako M., Shinomura T., Ujita M., Ito K., Kimata K.;
RT "Expression of PG-M(V3), an alternatively spliced form of PG-M
RL without a chondroitin sulfate attachment in region in mouse and human
RN tissues.";
RP [7]
RX J. Biol. Chem. 270:3914-3918(1995).
RP SEQUENCE OF 3333-3396 FROM N.A. (ISOFORM VINT).
RX TISSUE=Aortic smooth muscle;
RA Lemire J.M., Braun K.R., Maurel P., Kaplan E.D., Schwartz S.M.,
RT Wight T.N.;
RL "Versican/PG-M isoforms in vascular smooth muscle cells.";
RN [8]
RX Arterioscler. Thromb. Vasc. Biol. 19:1630-1639(1999).
RP PARTIAL SEQUENCE.
RX TISSUE=Brain;
RA Perides G., Lane W.S., Andrews D., Dahl D., Bignami A.;
RT "Isolation and partial characterization of a glial
RL hyaluronate-binding protein.";
RN [9]
RX J. Biol. Chem. 264:5981-5987(1989).
RP TISSUE SPECIFICITY OF ISOFORMS.
RX MEDLINE=96213482; PubMed=8627343;
```


RA Paulus W., Baur I., Dours-Zimmermann M.T., Zimmermann D.R.;
RT "Differential expression of versican isoforms in brain tumors.";
RL J. Neuropathol. Exp. Neurol. 55:528-533(1996).
CC -!- FUNCTION: May play a role in intracellular signaling and in
CC connecting cells with the extracellular matrix. May take part in
CC the regulation of cell motility, growth and differentiation. Binds
CC hyaluronic acid.
CC -!- SUBUNIT: Interacts with FBLN1 (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=5;
CC Comment=Additional isoforms seem to exist;
CC Name=V0;
CC IsoId=P13611-1; Sequence=Displayed;
CC Name=V1;
CC IsoId=P13611-2; Sequence=VSP_003082, VSP_003083;
CC Name=V2;
CC IsoId=P13611-3; Sequence=VSP_003084;
CC Name=V3;
CC IsoId=P13611-4; Sequence=VSP_003082, VSP_003085;
CC Name=Vint;
CC IsoId=P13611-5; Sequence=VSP_003086;
CC -!- TISSUE SPECIFICITY: Cerebral white matter. V0 and V1 is expressed
CC in normal brain, gliomas, medulloblastomas, schwannomas,
CC neurofibromas, and meningiomas; V2 is restricted to normal brain
CC and gliomas; V3 is found in all these tissues except
CC medulloblastomas.
CC -!- DEVELOPMENTAL STAGE: Disappears after the cartilage development.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -!- SIMILARITY: Contains 2 link domains.
CC -!- SIMILARITY: Contains 2 EGF-like domains.
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC -!- SIMILARITY: Contains 1 Sushi (SCR) domain.
CC -!- SIMILARITY: BELONGS TO THE AGGRECAN/VERSICAN PROTEOGLYCAN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U16306; AAG5018.1; -.
DR EMBL; X15998; CAA34128.1; -.
DR EMBL; S52488; AAB24878.1; -.
DR EMBL; U26555; AAA67565.1; -.
DR EMBL; D32039; BAA06801.1; -.
DR EMBL; J02814; AAA36437.1; -.
DR EMBL; AF084545; AAD48545.1; -.
DR PIR; S06014; A60979.
DR HSSP; P01132; 1EGF.
DR Genew; HGNC:2464; CSPG2.
DR MIM; 118661; -.
DR GO; GO:0005578; C:extracellular matrix; TAS.
DR GO; GO:0005204; F:chondroitin sulfate proteoglycan; TAS.
DR GO; GO:0005540; F:hyaluronic acid binding activity; TAS.
DR GO; GO:0008037; P:cell recognition; TAS.
DR GO; GO:0007275; P:development; TAS.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR007110; Ig_Like.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR000538; Link.
DR InterPro; IPR000436; Sushi_SCR_CCP.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00059; lectin_c; 1.
DR Pfam; PF00084; sushi_1.
DR Pfam; PF00193; Xlink; 2.

DR ProDom; PD000918; Link; 2.
DR SMART; SM00032; CCP; 1.
DR SMART; SM00034; CLECT; 1.
DR SMART; SM00179; EGF_CA; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00445; LINK; 2.
DR PROSITE; PS00010; ASX_HYDROXYL; 1.
DR PROSITE; PS00615; C_TYPE_LLECTIN_1; 1.
DR PROSITE; PS00041; C_TYPE_LLECTIN_2; 1.
DR PROSITE; PS00022; EGF_1_2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01187; EGF_CA; 1.
DR PROSITE; PS00835; IG_LIKE; 1.
DR PROSITE; PS01241; LINK; 2.
KW Glycoprotein; Proteoglycan; Lectin; Extracellular matrix; Sushi;
KW Signal; Repeat; EGF-like domain; Calcium; Immunoglobulin domain;
KW Hyaluronic acid; Alternative splicing.
FT SIGNAL 1 20
FT CHAIN 21 3396
FT DOMAIN 21 146
FT DOMAIN 167 244
FT DOMAIN 265 346
FT DOMAIN 348 1335
FT DOMAIN 1336 3089
FT DOMAIN 3089 3125
FT DOMAIN 3127 3163
FT DOMAIN 3176 3290
FT DOMAIN 3295 3353
FT DISULFID 44 130
FT DISULFID 172 243
FT DISULFID 196 217
FT DISULFID 270 345
FT DISULFID 294 315
FT DISULFID 3093 3104
FT DISULFID 3098 3113
FT DISULFID 3115 3124
FT DISULFID 3131 3142
FT DISULFID 3136 3151
FT DISULFID 3153 3162
FT DISULFID 3169 3180
FT DISULFID 3197 3289
FT DISULFID 3265 3281
FT DISULFID 3296 3339
FT DISULFID 3325 3352
FT CARBOHYD 57 57
FT CARBOHYD 330 330
FT CARBOHYD 615 615
Query Match 15.8%; Score 174; DB 1; Length 3396;
Best Local Similarity 28.5%; Pred. No. 4.9e-07;
Matches 47; Conservative 23; Mismatches 55; Indels 40; Gaps 8;
QY 14 QRPCYKVIYFHDTSRRLNFEAEKACRRDGGQLVSIIESEDEQKLIKFIENLLPSDGF- 72
Db 3177 QGQCYK--YF---AHRRTWDAARECRRLQGAHLTSILSHEEQMFVNRV-----GHDIQ 3224
QY 73 WIGLRRREKQSNSTACQDLVATDGSISQFRNWWYVDEP-----SCGSEVCVVMYHQSAP 128
Db 3225 WIGL-----NDKMFEDHDFRTDGTSTLOYENWRPNQPDSPFSAGEDCVIWHENG-- 3274
QY 129 AGIGGPFMFQWDDRCNMKNFKICKYS-----DEKPAVPSREAGE 169
Db 3275 -----QWDDVPCNYHLTYTCKKGTGTVACGQPPVVENAKTFGK 3310
RESULT 11
PGCV_CHICK
ID_PGCV_CHICK STANDARD; PRT; 3562 AA.
AC Q90953; Q90945;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Versican core protein precursor (Large fibroblast proteoglycan)
DE (Chondroitin sulfate proteoglycan core protein 2) (PG-M).
GN CPSG2.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS V0 AND V1).
RC STRAIN=White leghorn; TISSUE=Limb bud;
RX MEDLINE=93300846; PubMed=8314802;
RA Shinomura T., Nishida Y., Ito K., Kimata K.;
RT "cDNA cloning of PG-M, a large chondroitin sulfate proteoglycan
RT expressed during chondrogenesis in chick limb buds. Alternative
RT spliced multiforms of PG-M and their relationships to versican.";
RL J. Biol. Chem. 268:14461-14469(1993).
CC -!- FUNCTION: May play a role in intercellular signaling and in
CC connecting cells with the extracellular matrix. May take part in
CC the regulation of cell motility, growth and differentiation. Binds
CC hyaluronic acid.
CC -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Comment=Additional isoforms seem to exist;
CC Name=V0;
CC IsoId=Q90953-1; Sequence=Displayed;
CC Name=V1;
CC IsoId=Q90953-2; Sequence=VSP_003093;
CC TISSUE SPECIFICITY: Prechondrogenic condensation area of
CC developing limb buds.
CC -!- DEVELOPMENTAL STAGE: Disappears after the cartilage development
CC (By similarity).
CC -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -!- SIMILARITY: Contains 2 link domains.
CC -!- SIMILARITY: Contains 2 EGF-like domains.
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC -!- SIMILARITY: Contains 1 Sushi (SCR) domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X60226; CAA42787.1; -.
DR EMBL; D13542; BAA02742.1; -.
DR PIR; A47171; A47171.
DR HSSP; P00740; 1EDM.
DR InterPro; IPR000152; Abs_hydroxyl.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF_Like.
DR InterPro; IPR007110; Ig_Like.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR000538; Link.
DR InterPro; IPR000436; Sushi_SCR_CCP.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00059; lectin_c; 1.
DR Pfam; PF00084; sushi; 1.
DR Pfam; PF00193; Xlink; 2.
DR PRINTS; PR01265; LINKMODULE.
DR ProDom; PD000918; Link; 2.
DR SMART; SM00032; CCP; 1.
DR SMART; SM00034; CLECT; 1.
DR SMART; SM00179; EGF_CA; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00445; LINK; 2.

DR PROSITE; PS00010; ASX_HYDROXYL; 1.
DR PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01187; EGF_CA; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR PROSITE; PS01241; LINK; 2.
KW Glycoprotein; Proteoglycan; Lectin; Extracellular matrix; Sushi;
KW Signal; Repeat; EGF-like domain; Calcium; Immunoglobulin domain;
KW Hyaluronic acid; Alternative splicing.
FT SIGNAL 1 26
FT CHAIN 27 3562
FT DOMAIN 27 143
FT DOMAIN 166 243
FT DOMAIN 264 345
FT DOMAIN 3254 3290
FT DOMAIN 3292 3328
FT DOMAIN 3341 3455
FT DOMAIN 3460 3518
FT DISULFID 44 129
FT DISULFID 171 242
FT DISULFID 195 216
FT DISULFID 269 344
FT DISULFID 293 314
FT DISULFID 3258 3269
FT DISULFID 3263 3278
FT DISULFID 3280 3289
FT DISULFID 3296 3307
FT DISULFID 3301 3316
FT DISULFID 3318 3327
FT DISULFID 3334 3345
FT DISULFID 3362 3454
FT DISULFID 3430 3446
FT DISULFID 3461 3504
FT DISULFID 3490 3517
FT CARBOHYD 163 163
FT CARBOHYD 235 235
FT CARBOHYD 329 329
FT CARBOHYD 529 529
FT CARBOHYD 709 709
FT CARBOHYD 948 948
FT CARBOHYD 1409 1409
FT CARBOHYD 1479 1479
FT CARBOHYD 1523 1523
FT CARBOHYD 1530 1530
FT CARBOHYD 1625 1625
FT CARBOHYD 1751 1751
FT CARBOHYD 1988 1988
FT CARBOHYD 2088 2088
FT CARBOHYD 2089 2089
FT CARBOHYD 2507 2507
FT CARBOHYD 2642 2642
FT CARBOHYD 2679 2679
FT CARBOHYD 2748 2748
FT CARBOHYD 2762 2762
FT CARBOHYD 3069 3069
FT CARBOHYD 3194 3194
FT CARBOHYD 3232 3232
FT CARBOHYD 3545 3545
FT VARSPLIC 485 1411
FT SQ SEQUENCE 3562 AA; 388078 MW; 9BC565E88C1602D2 CRC64;
/FTid=VSP_003093.
Query Match 15.5%; Score 171; DB 1; Length 3562;
Best Local Similarity 28.5%; Pred. No. 9.7e-07;
Matches 47; Conservative 23; Mismatches 55; Indels 40; Gaps 8;
QY 14 QRPCYKVIYFHDTSRRLNFEAEKACRRDGGQGLVSISEDEQKLEKFIENLLPSDGF- 72
Db 3342 QGQCYK--YF---AHRRTWDTAERECRLQGAHLTSILSHEEQVFNRI-----GHDIQ 3389
QY 73 WIGLRRREKQSNSTACQDLYAWTDGSIQFRNWTVDPE-----SCGSEVCVVMYHQPAP 128

GN AGC1.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Hering T.M., Kollar J., Huynh T.D.;
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 563-1056 FROM N.A.
RX MEDLINE=89380219; PubMed=2528543;
RA Antonsson P., Heinigaard D., Oldberg A.;
RT "The keratan sulfate-enriched region of bovine cartilage proteoglycan
RT consists of a consecutively repeated hexapeptide motif.";
RL J. Biol. Chem. 264:16170-16173(1989).
RN [3]
RP SEQUENCE OF 1609-2113 AND 2151-2364 FROM N.A.
RX MEDLINE=87270630; PubMed=3111460;
RA Oldberg A., Antonsson P., Heinigaard D.;
RT "The partial amino acid sequence of bovine cartilage proteoglycan,
RT deduced from a cDNA clone, contains numerous Ser-Gly sequences
RT arranged in homologous repeats.";
RL Biochem. J. 243:255-259(1987).
RN [4]
RP SEQUENCE OF 2114-2150 FROM N.A.
RC TISSUE=Cartilage;
RX MEDLINE=93352525; PubMed=8349621;
RA Fueleop C., Walcz E., Valyon M., Glant T.T.;
RT "Expression of alternatively spliced epidermal growth factor-like
RT domains in aggregates of different species. Evidence for a novel
RT module.";
RL J. Biol. Chem. 268:17377-17383(1993).
RN [5]
RP PARTIAL SEQUENCE.
RX MEDLINE=85027710; PubMed=6489519;
RA Perin J.-P., Bonnet F., Jolles J., Jolles P.;
RT "Sequence data concerning the protein core of the cartilage
RT proteoglycan monomers. Characterization of a sequence allowing the
RT synthesis of an oligonucleotide probe.";
RL FEBS Lett. 176:37-42(1984).
RN [6]
RP PARTIAL SEQUENCE.
RX MEDLINE=87005253; PubMed=3530809;
RA Perin J.P., Bonnet F., Jolles P.;
RT "Structural relationship between link proteins and proteoglycan
RT monomers.";
RL FEBS Lett. 206:73-77(1986).
CC -!- FUNCTION: THIS PROTEOGLYCAN IS A MAJOR COMPONENT OF EXTRACELLULAR
CC MATRIX OF CARTILAGINOUS TISSUES. A MAJOR FUNCTION OF THIS PROTEIN
CC IS TO RESIST COMPRESSION IN CARTILAGE. IT BINDS AVIDLY TO
CC HYALURONIC ACID VIA AN AMINO-TERMINAL GLOBULAR REGION. MAY PLAY A
CC REGULATORY ROLE IN THE MATRIX ASSEMBLY OF THE CARTILAGE.
CC -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix (By
CC similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=P13608-1; Sequence=Displayed;
CC Name=2;
CC IsoId=P13608-2; Sequence=VSP_003072;
CC -!- DOMAIN: TWO GLOBULAR DOMAINS, G1 AND G2, COMPRISE THE AMINO
CC TERMINUS OF THE PROTEOGLYCAN, WHILE ANOTHER GLOBULAR REGION, G3,
CC MAKES UP THE C-TERMINUS. G1 CONTAINS LINK DOMAINS AND THUS
CC CONSISTS OF THREE DISULFIDE-BONDED LOOP STRUCTURES DESIGNATED AS
CC THE A, B, B' MOTIFS. G2 IS SIMILAR TO G1. THE KERATAN SULFATE (KS)
CC AND THE CHONDROITIN SULFATE (CS) ATTACHMENT DOMAINS LIE BETWEEN G2
CC AND G3.
CC -!- PTM: CONTAINS MOSTLY CHONDROITIN SULFATE, BUT ALSO N-LINKED AND O-
CC LINKED (ABOUT 40) OLIGOSACCHARIDES.
CC -!- PTM: THE KERATAN SULFATE CONTENTS DIFFER CONSIDERABLY BETWEEN
CC ADULT AND FETAL BOVINE PROTEOGLYCANS.

CC -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -!- SIMILARITY: Contains 4 link domains.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC -!- SIMILARITY: Contains 1 Sushi (SCR) domain.
CC -!- SIMILARITY: BELONGS TO THE AGGRECAN/VERSICAN PROTEOGLYCAN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U76615; AAB38524.1; -;
DR EMBL; L07053; -; NOT_ANNOTATED_CDS.
DR PIR; A34234; A39808.
DR PIR; T42630; T42630.
DR HSSP; P08709; 1BF9.
DR InterPro; IPR002353; AntifreezeII.
DR InterPro; IPR000152; Asx hydroxyl.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR000538; Link.
DR InterPro; IPR003324; SGXSG.
DR InterPro; IPR000436; Sushi_SCR_CCP.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF00059; lectin_c; 1.
DR Pfam; PF02339; SGXSG; 61.
DR Pfam; PF00084; sushi; 1.
DR Pfam; PF00193; Xlink; 4.
DR PRINTS; PR01265; LINKMODULE.
DR PRINTS; PR00356; ANTIFREEZEII.
DR ProDom; PD000918; Link; 4.
DR SMART; SM00032; CCP; 1.
DR SMART; SM00034; CLECT; 1.
DR SMART; SM00179; EGF_CA; 1.
DR SMART; SM00445; LINK; 4.
DR PROSITE; PS00010; ASX HYDROXYL; 1.
DR PROSITE; PS00615; C_TYPE_LLECTIN_1; 1.
DR PROSITE; PS50041; C_TYPE_LLECTIN_2; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01187; EGF_CA; 1.
DR PROSITE; PS50835; IG LIKE; 1.
DR PROSITE; PS00290; IG_MHC; 1.
DR PROSITE; PS01241; LINK; 4.
KW Glycoprotein; Cartilage; Proteoglycan; Lectin; Signal; Sushi;
KW EGF-like domain; Calcium; Alternative splicing; Repeat;
KW Immunoglobulin domain.
FT SIGNAL 1 16 POTENTIAL.
FT CHAIN 17 2364 AGGRECAN CORE PROTEIN.
FT DOMAIN 25 147 IG-LIKE V-TYPE.
FT DOMAIN 170 247 LINK 1.
FT DOMAIN 268 349 LINK 2.
FT DOMAIN 504 581 LINK 3.
FT DOMAIN 602 683 LINK 4.
FT DOMAIN 774 907 23 X 6 AA APPROXIMATE TANDEM REPEATS OF
FT E-[EK]-P-F-P-S.
FT CS-2.
FT EGF-LIKE, CALCIUM-BINDING (POTENTIAL).
FT G3.
FT C-TYPE LECTIN.
FT SUSHI.
FT BY SIMILARITY.
FT BY SIMILARITY.
FT BY SIMILARITY.
FT BY SIMILARITY.
FT DISULFID 51 133
FT DISULFID 175 246
FT DISULFID 199 220
FT DISULFID 273 348

FT DISULFID 297 318 BY SIMILARITY.
FT DISULFID 509 580 BY SIMILARITY.
FT DISULFID 533 554 BY SIMILARITY.
FT DISULFID 607 682 BY SIMILARITY.
FT DISULFID 631 652 BY SIMILARITY.
FT DISULFID 2117 2128 BY SIMILARITY.
FT DISULFID 2182 2274 BY SIMILARITY.
FT DISULFID 2250 2266 BY SIMILARITY.
FT DISULFID 2281 2324 BY SIMILARITY.
FT DISULFID 2310 2337 BY SIMILARITY.
FT CARBOHYD 126 126 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 239 239 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 333 333 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 387 387 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 611 611 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 667 667 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 2114 2150 Missing (in isoform 2).
FT /FTid=VSP 003072.
SQ SEQUENCE 2364 AA; 246359 MW; 6FF83763420C3D4C CRC64;
Query Match 14.4%; Score 158.5; DB 1; Length 2364;
Best Local Similarity 26.8%; Pred. No. 7.7e-06;
Matches 48; Conservative 29; Mismatches 59; Indels 43; Gaps 10;
Oy 6 QPVCRRG---TQPCYKVIYFHTTSRRLNFEAEACRRDGGQLVSIIESEDEKLIKFI 62
Db 2151 QKLCBEGWTKFGHCYR--HFPD---RATVDAESQCRKQSHLSSIVTPEEQ----EFV 2201
Oy 63 ENLLPSDGDGF-WIGLRRRREKQSNSTACQDLYAWTDGSGISOFRNWYVDEP----SCGSEV 117
Db 2202 NN---NAQDYQWIGL-----NDKTIEGDFRWSGDGHSLOFENWRPNQPDNFFATGEDC 2250
Oy 118 CVVMYHQPAPAGIGGYPWFQNDRCNMKNFICKYS-----DEKPAVPSREAEGEETE 172
Db 2251 VVMWHEKG-----EWNDVPCNYQLPFTCKKGTVACGEPVVEHARIFGQKXD 2298
RESULT 14
PGCA_HUMAN
ID PGCA_HUMAN STANDARD; PRT; 2415 AA.
AC P16112; Q13650;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Aggrecan core protein precursor (Cartilage-specific proteoglycan core protein)
DE protein) (CSPCP) (Chondroitin sulfate proteoglycan core protein 1).
GN AGC1 OR CSPG1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 3).
RC TISSUE=Chondrocytes;
RX MEDLINE=91093289; PubMed=1985970;
RA Doege K.J., Sasaki M., Kimura T., Yamada Y.;
RT "Complete coding sequence and deduced primary structure of the human
RT cartilage large aggregating proteoglycan, aggrecan. Human-specific
RT repeats, and additional alternatively spliced forms.";
RL J. Biol. Chem. 266:894-902(1991).
RN [2]
RP SEQUENCE OF 1778-2415 FROM N.A. (ISOFORM 2).
RC TISSUE=Chondrocytes;
RA Dudhia J., Hardingham T.E.;
RL Submitted (JAN-1990) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 1936-2415 FROM N.A. (ISOFORM 1).
RX MEDLINE=89380154; PubMed=2789216;
RA Baldwin C.T., Reginato A.M., Prockop D.J.;
RT "A new epidermal growth factor-like domain in the human core protein
RT for the large cartilage-specific proteoglycan. Evidence for
RT alternative splicing of the domain.";
RL J. Biol. Chem. 264:15747-15750(1989).

RN [4]
RP SEQUENCE OF 764-864 FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=95128522; PubMed=7827755;
RA Barry F.P., Neame P.J., Sasse J., Pearson D.;
RT "Length variation in the keratan sulfate domain of mammalian
RT aggrecan.";
RL Matrix Biol. 14:323-328(1994).
CC -1- FUNCTION: THIS PROTEOGLYCAN IS A MAJOR COMPONENT OF EXTRACELLULAR
CC MATRIX OF CARTILAGENOUS TISSUES. A MAJOR FUNCTION OF THIS PROTEIN
CC IS TO RESIST COMPRESSION IN CARTILAGE. IT BINDS AVIDLY TO
CC HYALURONIC ACID VIA AN AMINO-TERMINAL GLOBULAR REGION.
CC -1- SUBUNIT: Interacts with FBLN1 (By similarity).
CC -1- SUBCELLULAR LOCATION: Secreted; extracellular matrix (By
CC similarity).
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Comment=Additional isoforms seem to exist;
CC Name=1;
CC IsoId=P16112-1; Sequence=Displayed;
CC Name=2;
CC IsoId=P16112-2; Sequence=VSP_003074;
CC Name=3;
CC IsoId=P16112-3; Sequence=VSP_003074, VSP_003075;
CC -1- DOMAIN: TWO GLOBULAR DOMAINS, G1 AND G2, COMPRISE THE AMINO
CC TERMINUS OF THE PROTEOGLYCAN, WHILE ANOTHER GLOBULAR REGION, G3,
CC MAKES UP THE COOH TERMINUS. G1 CONTAINS LINK DOMAINS AND THUS
CC CONSISTS OF THREE DISULFIDE-BONDED LOOP STRUCTURES DESIGNATED AS
CC THE A, B, B' MOTIFS. G2 IS SIMILAR TO G1. THE KERATAN SULFATE (KS)
CC AND THE CHONDROITIN SULFATE (CS) ATTACHMENT DOMAINS LIE BETWEEN G2
CC AND G3.
CC -1- PTM: CONTAINS MOSTLY CHONDROITIN SULFATE, BUT ALSO KERATAN SULFATE
CC CHAINS, N-LINKED AND O-LINKED OLIGOSACCHARIDES.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -1- SIMILARITY: Contains 4 link domains.
CC -1- SIMILARITY: Contains 1 EGF-like domain.
CC -1- SIMILARITY: Contains 1 C-type lectin family domain.
CC -1- SIMILARITY: Contains 1 Sushi (SCR) domain.
CC -1- SIMILARITY: BELONGS TO THE AGGRECAN/VERSICAN PROTEOGLYCAN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M55172; AAA62824.1; -.
CC EMBL; J05062; AAA35726.1; -.
CC EMBL; X17406; CAA35463.1; -.
CC EMBL; S74659; AAC60643.2; -.
CC PIR; A39086; A39086.
CC HSSP; P98066; 1TSG.
CC Genew; HGNC:319; AGC1.
CC MIM; 155760; -.
CC GO; GO:0005204; F:chondroitin sulfate proteoglycan; TAS.
CC InterPro; IPR002353; AntifreezeII.
CC InterPro; IPR006209; EGF like.
CC InterPro; IPR006210; IEGF.
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR003599; Ig.
CC InterPro; IPR003006; Ig_MHC.
CC InterPro; IPR001304; Lectin_C.
CC InterPro; IPR000538; Link.
CC InterPro; IPR003324; SGXXSG.
CC InterPro; IPR000436; Sushi_SCR_CCP.
CC Pfam; PF00008; EGF; 1.
CC Pfam; PF00047; ig; 1.
CC Pfam; PF00059; lectin_c; 1.
CC Pfam; PF02339; SGXXSG; 71.
CC Pfam; PF00084; sushi; 1.
CC Pfam; PF00193; Xlink; 4.

CC AND THE CHONDROITIN SULFATE (CS) ATTACHMENT DOMAINS LIE BETWEEN G2
CC AND G3.
CC -!- PTM: CONTAINS MOSTLY CHONDROITIN SULFATE, BUT ALSO KERATAN SULFATE
CC CHAINS, N-LINKED AND O-LINKED OLIGOSACCHARIDES (BY SIMILARITY).
CC -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -!- SIMILARITY: Contains 4 link domains.
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC -!- SIMILARITY: Contains 1 Sushi (SCR) domain.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC -!- SIMILARITY: BELONGS TO THE AGGRECAN/VERSICAN PROTEOGLYCAN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U65989; AAC06238.2; -
CC EMBL; S74662; AAC60527.1; -
CC EMBL; L07054; -; NOT_ANNOTATED_CDS.
CC PIR; I46998; I46998.
CC HSSP; P08709; 1BF9.
CC InterPro; IPR000152; Asx_hydroxyl.
CC InterPro; IPR000742; EGF_2.
CC InterPro; IPR001881; EGF_Ca.
CC InterPro; IPR006209; EGF_like.
CC InterPro; IPR007110; Ig_Like.
CC InterPro; IPR003006; Ig_MHC.
CC InterPro; IPR001304; Lectin_C.
CC InterPro; IPR000538; Linkin_C.
CC InterPro; IPR003324; SGXXSG.
CC InterPro; IPR000436; Sushi_SCR_CCP.
CC Pfam; PF00008; EGF; 1.
CC Pfam; PF00047; ig; 1.
CC Pfam; PF00059; lectin_c; 1.
CC Pfam; PF02339; SGXXSG; 66.
CC Pfam; PF00084; sushi; 1.
CC Pfam; PF00193; Xlink; 4.
CC PRINTS; PR01265; LINKMODULE.
CC PRINTS; PR00356; ANTIFREEZEII.
CC ProDom; PD000918; Link; 4.
CC SMART; SM00032; CCP; 1.
CC SMART; SM00034; CLECT; 1.
CC SMART; SM00179; EGF_CA; 1.
CC SMART; SM00445; LINK; 4.
CC PROSITE; PS00010; ASX_HYDROXYL; 1.
CC PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
CC PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
CC PROSITE; PS00022; EGF_1; 1.
CC PROSITE; PS01187; EGF_CA; 1.
CC PROSITE; PS50835; IG_LIKE; 1.
CC PROSITE; PS00290; IG_MHC; 1.
CC PROSITE; PS01241; LINK; 4.
KW Glycoprotein; Cartilage; Proteoglycan; Lectin; Signal; Sushi;
KW EGF-like domain; Repeat; Immunoglobulin domain.
FT SIGNAL 1 16 POTENTIAL.
FT CHAIN 17 2333 AGGRECAN CORE PROTEIN.
FT DOMAIN 34 147 IG-LIKE V-TYPE.
FT DOMAIN 170 247 LINK 1.
FT DOMAIN 268 349 LINK 2.
FT DOMAIN 513 590 LINK 3.
FT DOMAIN 611 692 LINK 4.
FT DOMAIN 2081 2117 EGF-LIKE, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 2130 2245 C-TYPE LECTIN.
FT DOMAIN 2249 2307 SUSHI.
FT DOMAIN 48 140 G1-A.
FT DOMAIN 152 247 G1-B.
FT DOMAIN 253 349 G1-B'.
FT DOMAIN 495 589 G2-B.
FT DOMAIN 596 691 G2-B'.
FT DOMAIN 694 816 KS.

FT	DOMAIN	819	1394	CS-1.
FT	DOMAIN	1395	2079	CS-2.
FT	DOMAIN	2080	2333	G3.
FT	DISULFID	51	133	BY SIMILARITY.
FT	DISULFID	175	246	BY SIMILARITY.
FT	DISULFID	199	220	BY SIMILARITY.
FT	DISULFID	273	348	BY SIMILARITY.
FT	DISULFID	297	318	BY SIMILARITY.
FT	DISULFID	518	589	BY SIMILARITY.
FT	DISULFID	542	563	BY SIMILARITY.
FT	DISULFID	616	691	BY SIMILARITY.
FT	DISULFID	640	661	BY SIMILARITY.
FT	DISULFID	2085	2096	BY SIMILARITY.
FT	DISULFID	2090	2105	BY SIMILARITY.
FT	DISULFID	2107	2116	BY SIMILARITY.
FT	DISULFID	2123	2134	BY SIMILARITY.
FT	DISULFID	2151	2243	BY SIMILARITY.
FT	DISULFID	2219	2235	BY SIMILARITY.
FT	DISULFID	2250	2293	BY SIMILARITY.
FT	DISULFID	2279	2306	BY SIMILARITY.
FT	CARBOHYD	126	126	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	239	239	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	333	333	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	387	387	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	444	444	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	620	620	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	676	676	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	747	747	N-LINKED (GLCNAC. .) (POTENTIAL).
SQ	SEQUENCE	2333 AA;	240573 MW;	8B9ED78F3508B596 CRC64;

Query Match 14.1%; Score 155.5; DB 1; Length 2333;
Best Local Similarity 28.5%; Pred. No. 1.4e-05;
Matches 51; Conservative 24; Mismatches 61; Indels 43; Gaps 11;

QY	6	QPVCRRG	---TQRPCYKVIYFHDTSRRRLNPFEEAKEACRRDGGQLVSI	SEDEQKLI	EKF 62
Db	2120	QELCEEGWTKFQGH	CYR--YFPD---RESWVDAESRCRAQQSHLSSIVTPEEQ	----EFV	2170
QY	63	ENLLPSDGD	DF-WIGLRRRREKQSNSTACQDLYAWTDGSI	SQFRN	WYVDEPS 118
Db	2171	NN---NAQDYQW	IGL-----NDRTIEGDFRWSGDGHS	LQFENWRPNQPDN	FFVSGEDC 2219
QY	119	VVM-YHQPSA	PAGIGGYPYMFQWDDRCNMKNFICKYS	----DEKPAVPS	REAEGEETE 172
Db	2220	VVMIWHEKG	-----EWDVPCNYLPFTCKKGT	VACGDPVVEHART	FGQKGD 2267

Search completed: December 22, 2003, 16:10:59
Job time : 7.46605 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 22, 2003, 16:06:34 ; Search time 9.86924 Seconds
(without alignments)
1978.090 Million cell updates/sec

Title: US-09-887-855-2_COPY_25_227
Perfect score: 1100
Sequence: 1 RLLSGQPVCRGGTQRPCYKV.....EEDAKKTPKESREAAALNLAY 203

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 76:*
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	182	16.5	1456	1 A36563	mannose receptor p
2	178.5	16.2	1455	1 A48925	mannose receptor p
3	177	16.1	1268	2 S52781	neurocan - mouse
4	174.5	15.9	1643	2 T14274	versican precursor
5	174.5	15.9	3381	2 T42389	versican precursor
6	174	15.8	1257	2 S28764	neurocan precursor
7	174	15.8	2397	1 A55535	versican precursor
8	174	15.8	2409	1 A60979	versican precursor
9	171	15.5	3562	2 A47171	chondroitin sulfat
10	170.5	15.5	1479	2 T42710	mannose receptor,
11	158.5	14.4	1340	2 A39808	proteoglycan core
12	158.5	14.4	2327	2 T42630	aggreacan - bovine
13	158.5	14.4	2415	1 A39086	aggreacan precursor
14	154.5	14.0	612	2 B42755	E-selectin precurs
15	153.5	14.0	2124	2 A28452	proteoglycan core
16	152	13.8	912	2 A54423	brevican precursor
17	149.5	13.6	459	2 T24425	hypothetical prote
18	149	13.5	321	1 LNHUER	IgE Fc receptor II
19	148.5	13.5	330	2 T46256	brevican - human (
20	148	13.5	253	2 E89130	protein F52E1.2 [i
21	147.5	13.4	2132	1 A55182	aggreacan precursor
22	147	13.4	883	2 S57653	brevican precursor
23	146.5	13.3	162	1 LNRCl	lectin BRA3-1 prec
24	146	13.3	2109	1 I50421	aggreacan precursor
25	145.5	13.2	742	2 JC7595	scavenger receptor
26	145	13.2	883	2 S49126	brevican precursor
27	144.5	13.1	173	2 S10548	lectin - barnacle
28	144.5	13.1	372	2 S23936	L-selectin precurs
29	144.5	13.1	404	2 A46274	HIV gp120-binding

30	143.5	13.0	129	2 JC4329	coagulation factor
31	143.5	13.0	372	1 A32375	L-selectin precurs
32	143.5	13.0	463	2 T26655	hypothetical prote
33	142.5	13.0	131	2 JC5058	bitisacetin alpha c
34	142	12.9	1487	2 S48719	phospholipase-A(2)
35	141.5	12.9	331	1 LNMSEK	IgE Fc receptor, 1
36	140.5	12.8	162	1 LNRCl	lectin BRA3-2 prec
37	140	12.7	370	2 S22124	L-selectin precurs
38	139	12.6	248	1 LNHUPS	pulmonary surfacta
39	139	12.6	248	1 LNHUP6	pulmonary surfacta
40	139	12.6	248	1 LNHUP1	pulmonary surfacta
41	139	12.6	283	1 LNFHLS	lectin precursor -
42	138.5	12.6	152	2 JC4690	coagulation factor
43	138.5	12.6	202	2 JC4031	tetranectin precu
44	138	12.5	280	2 T29200	hypothetical prote
45	137.5	12.5	309	1 S34198	IgE Fc receptor II

ALIGNMENTS

RESULT 1

A36563

mannose receptor precursor - human

C;Species: Homo sapiens (man)

C;Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text change 10-Sep-1999

C;Accession: A36563; A60926; A44255; B44255; C44255; D44255; E44255; F44255; G44255; H44255; I44255; J44255; K44255; L44255; M44255; N44255; O44255; P44255; Q44255; R44255; S44255; T44255; U44255; V44255; W44255; X44255; Y44255; Z44255; AA44255; AB44255; AC44255; AD44255; AE44255; AF44255; AG44255; AH44255; AI44255; AJ44255; AK44255; AL44255; AM44255; AN44255; AO44255; AP44255; AQ44255; AR44255; AS44255; AT44255; AU44255; AV44255; AW44255; AX44255; AY44255; AZ44255; BA44255; BB44255; BC44255; BD44255; BE44255; BF44255; BG44255; BH44255; BI44255; BJ44255; BK44255; BL44255; BM44255; BN44255; BO44255; BP44255; BQ44255; BR44255; BS44255; BT44255; BU44255; BV44255; BW44255; BX44255; BY44255; BZ44255; CA44255; CB44255; CC44255; CD44255; CE44255; CF44255; CG44255; CH44255; CI44255; CJ44255; CK44255; CL44255; CM44255; CN44255; CO44255; CP44255; CQ44255; CR44255; CS44255; CT44255; CU44255; CV44255; CW44255; CX44255; CY44255; CZ44255; DA44255; DB44255; DC44255; DD44255; DE44255; DF44255; DG44255; DH44255; DI44255; DJ44255; DK44255; DL44255; DM44255; DN44255; DO44255; DP44255; DQ44255; DR44255; DS44255; DT44255; DU44255; DV44255; DW44255; DX44255; DY44255; DZ44255; EA44255; EB44255; EC44255; ED44255; EE44255; EF44255; EG44255; EH44255; EI44255; EJ44255; EK44255; EL44255; EM44255; EN44255; EO44255; EP44255; EQ44255; ER44255; ES44255; ET44255; EU44255; EV44255; EW44255; EX44255; EY44255; EZ44255; FA44255; FB44255; FC44255; FD44255; FE44255; FF44255; FG44255; FH44255; FI44255; FJ44255; FK44255; FL44255; FM44255; FN44255; FO44255; FP44255; FQ44255; FR44255; FS44255; FT44255; FU44255; FV44255; FW44255; FX44255; FY44255; FZ44255; GA44255; GB44255; GC44255; GD44255; GE44255; GF44255; GG44255; GH44255; GI44255; GJ44255; GK44255; GL44255; GM44255; GN44255; GO44255; GP44255; GQ44255; GR44255; GS44255; GT44255; GU44255; GV44255; GW44255; GX44255; GY44255; GZ44255; HA44255; HB44255; HC44255; HD44255; HE44255; HF44255; HG44255; HH44255; HI44255; HJ44255; HK44255; HL44255; HM44255; HN44255; HO44255; HP44255; HQ44255; HR44255; HS44255; HT44255; HU44255; HV44255; HW44255; HX44255; HY44255; HZ44255; IA44255; IB44255; IC44255; ID44255; IE44255; IF44255; IG44255; IH44255; II44255; IJ44255; IK44255; IL44255; IM44255; IN44255; IO44255; IP44255; IQ44255; IR44255; IS44255; IT44255; IU44255; IV44255; IW44255; IX44255; IY44255; IZ44255; JA44255; JB44255; JC44255; JD44255; JE44255; JF44255; JG44255; JH44255; JI44255; JJ44255; JK44255; JL44255; JM44255; JN44255; JO44255; JP44255; JQ44255; JR44255; JS44255; JT44255; JU44255; JV44255; JW44255; JX44255; JY44255; JZ44255; KA44255; KB44255; KC44255; KD44255; KE44255; KF44255; KG44255; KH44255; KI44255; KL44255; KM44255; KN44255; KO44255; KP44255; KQ44255; KR44255; KS44255; KT44255; KU44255; KV44255; KW44255; KX44255; KY44255; KZ44255; LA44255; LB44255; LC44255; LD44255; LE44255; LF44255; LG44255; LH44255; LI44255; LJ44255; LK44255; LL44255; LM44255; LN44255; LO44255; LP44255; LQ44255; LR44255; LS44255; LT44255; LU44255; LV44255; LW44255; LX44255; LY44255; LZ44255; MA44255; MB44255; MC44255; MD44255; ME44255; MF44255; MG44255; MH44255; MI44255; MJ44255; MK44255; ML44255; MN44255; MO44255; MP44255; MQ44255; MR44255; MS44255; MT44255; MU44255; MV44255; MW44255; MX44255; MY44255; MZ44255; NA44255; NB44255; NC44255; ND44255; NE44255; NF44255; NG44255; NH44255; NI44255; NJ44255; NK44255; NL44255; NM44255; NO44255; NP44255; NQ44255; NR44255; NS44255; NT44255; NU44255; NV44255; NW44255; NX44255; NY44255; NZ44255; OA44255; OB44255; OC44255; OD44255; OE44255; OF44255; OG44255; OH44255; OI44255; OJ44255; OK44255; OL44255; OM44255; ON44255; OO44255; OP44255; OQ44255; OR44255; OS44255; OT44255; OU44255; OV44255; OW44255; OX44255; OY44255; OZ44255; PA44255; PB44255; PC44255; PD44255; PE44255; PF44255; PG44255; PH44255; PI44255; PJ44255; PK44255; PL44255; PM44255; PN44255; PO44255; PP44255; PQ44255; PR44255; PS44255; PT44255; PU44255; PV44255; PW44255; PX44255; PY44255; PZ44255; QA44255; QB44255; QC44255; QD44255; QE44255; QF44255; QG44255; QH44255; QI44255; QJ44255; QK44255; QL44255; QM44255; QN44255; QO44255; QP44255; QQ44255; QR44255; QS44255; QT44255; QU44255; QV44255; QW44255; QX44255; QY44255; QZ44255; RA44255; RB44255; RC44255; RD44255; RE44255; RF44255; RG44255; RH44255; RI44255; RJ44255; RK44255; RL44255; RM44255; RN44255; RO44255; RP44255; RQ44255; RR44255; RS44255; RT44255; RU44255; RV44255; RW44255; RX44255; RY44255; RZ44255; SA44255; SB44255; SC44255; SD44255; SE44255; SF44255; SG44255; SH44255; SI44255; SJ44255; SK44255; SL44255; SM44255; SN44255; SO44255; SP44255; SQ44255; SR44255; SS44255; ST44255; SU44255; SV44255; SW44255; SX44255; SY44255; SZ44255; TA44255; TB44255; TC44255; TD44255; TE44255; TF44255; TG44255; TH44255; TI44255; TJ44255; TK44255; TL44255; TM44255; TN44255; TO44255; TP44255; TQ44255; TR44255; TS44255; TT44255; TU44255; TV44255; TW44255; TX44255; TY44255; TZ44255; UA44255; UB44255; UC44255; UD44255; UE44255; UF44255; UG44255; UH44255; UI44255; UJ44255; UK44255; UL44255; UM44255; UN44255; UO44255; UP44255; UQ44255; UR44255; US44255; UT44255; UU44255; UV44255; UW44255; UX44255; UY44255; UZ44255; VA44255; VB44255; VC44255; VD44255; VE44255; VF44255; VG44255; VH44255; VI44255; VJ44255; VK44255; VL44255; VM44255; VN44255; VO44255; VP44255; VQ44255; VR44255; VS44255; VT44255; VU44255; VV44255; VW44255; VX44255; VY44255; VZ44255; WA44255; WB44255; WC44255; WD44255; WE44255; WF44255; WG44255; WH44255; WI44255; WJ44255; WK44255; WL44255; WM44255; WN44255; WO44255; WP44255; WQ44255; WR44255; WS44255; WT44255; WU44255; WV44255; WW44255; WX44255; WY44255; WZ44255; XA44255; XB44255; XC44255; XD44255; XE44255; XF44255; XG44255; XH44255; XI44255; XJ44255; XK44255; XL44255; XM44255; XN44255; XO44255; XP44255; XQ44255; XR44255; XS44255; XT44255; XU44255; XV44255; XW44255; XX44255; XY44255; XZ44255; YA44255; YB44255; YC44255; YD44255; YE44255; YF44255; YG44255; YH44255; YI44255; YJ44255; YK44255; YL44255; YM44255; YN44255; YO44255; YP44255; YQ44255; YR44255; YS44255; YT44255; YU44255; YV44255; YW44255; YX44255; YY44255; YZ44255; ZA44255; ZB44255; ZC44255; ZD44255; ZE44255; ZF44255; ZG44255; ZH44255; ZI44255; ZJ44255; ZK44255; ZL44255; ZM44255; ZN44255; ZO44255; ZP44255; ZQ44255; ZR44255; ZS44255; ZT44255; ZU44255; ZV44255; ZW44255; ZX44255; ZY44255; ZZ44255;

J. Biol. Chem. 265, 12156-12162, 1990

A;Title: Primary structure of the mannose receptor contains multiple motifs resembling

A;Reference number: A36563; MUID:90324192; PMID:2373685

A;Accession: A36563

A;Molecule type: mRNA

A;Residues: 1-1456 <TAY>

A;Cross-references: GB:J05550; NID:gl88675; PIDN:AAA59868.1; PID:gl88676

A;Note: parts of this sequence, including the amino end of the mature protein, were cor

R;Ezekowitz, R.A.B.; Sastry, K.; Bailly, P.; Warner, A.

J. Exp. Med. 172, 1785-1794, 1990

A;Title: Molecular characterization of the human macrophage mannose receptor: demonstr

A;Reference number: A60926; MUID:91079783; PMID:2258707

A;Accession: A60926

A;Status: nucleic acid sequence not shown

A;Molecule type: mRNA

A;Residues: 1-1333, 'T', 1335-1456 <EZE>

A;Cross-references: GB:X55635

A;Note: translation of the nucleotide sequence is incomplete

A;Note: in the authors' translation additional residues Pro-Glu-Ile are shown after 49;

R;Kim, S.J.; Ruiz, N.; Bezouska, K.; Drickamer, K.

Genomics 14, 721-727, 1992

A;Title: Organization of the gene encoding the human macrophage mannose receptor (MRC1)

A;Reference number: A44255; MUID:93052405; PMID:1294118

A;Accession: A44255

A;Status: preliminary; nucleic acid sequence not shown; not compared with conceptual t

A;Molecule type: DNA

A;Residues: 155-233, 'KSAL', 238-283; 346-428; 492-569; 631-714, 716-719; 783-820, 'N', 822-865,

A;Note: sequence extracted from NCBI backbone (NCBIP:118415, NCBIP:118421, NCBIP:118428

C;Genetics:

A;Gene: GDB:MRC1

A;Cross-references: GDB:133759; OMIM:153618

A;Map position: 10p13-10p13

C;Superfamily: phospholipase A2 receptor; C-type lectin homology; fibronectin type II

F;Keywords: duplication; lectin; tandem repeat; transmembrane protein

F;1-18/Domain: signal sequence #status predicted <SIG>

F;168-209/Domain: fibronectin type II repeat homology <2F1>

F;223-340/Domain: C-type lectin homology <LCH1>

F;362-486/Domain: C-type lectin homology <LCH2>

F;945-1079/Domain: C-type lectin homology <LCH3>

Query Match 16.5%; Score 182; DB 1; Length 1456;

Best Local Similarity 25.5%; Pred. No. 8.2e-08;

Matches 52; Conservative 37; Mismatches 61; Indels 54; Gaps 8;

QY

18 YKVIYFHTSRLNFEEAKEACRRDGGQLVSISEDEQKLEKFIENLLPSDGFWIGLR 77

Db 807 YKDYQYFSKETMDNARAFKRNFGDLVSIQSESEKFLWKYV-NRNDQAQSAFYIGLL 865
QY 78 RREEKQSNSTACQDLYAWTDGSIQFRNWWYVDEPSCGS--EVCVVMYHOPSAPAGIGGPY 135
Db 866 ISLDKK-----FAWMDGSKVDYVSWATGEPNPFANEDENCVTMY-----SNSGF---- 908
QY 136 MFQWNDRCNMKNPFICKYSDEK---PAVPSRAEAGEETELTTPVLPEETQE----- 184
Db 909 ---WNDINCGYPNPFICQRHNSINATTVM-----TPMSVPSCKEGWNFYSN 954
QY 185 -----EDAKKTFKESREAAAL 199
Db 955 KCFKIFGFMEERKNWQEARACI 978

RESULT 2
A48925
mannose receptor precursor, macrophage - mouse
C;Species: Mus musculus (house mouse)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A48925; S21320; PC2245
R;Harris, N.; Super, M.; Rits, M.; Chang, G.; Ezekowitz, R.A.
Blood 80, 2363-2373, 1992
A;Title: Characterization of the murine macrophage mannose receptor: demonstration that
A;Reference number: A48925; MUID:93043353; PMID:1421407
A;Accession: A48925
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-1455 <HAR>
A;Experimental source: peritoneal macrophage
A;Note: sequence extracted from NCBI backbone (NCBIP:118733)
R;Harris, N.; Super, M.; Rits, M.; Chang, G.; Ezekowitz, R.B.
submitted to the EMBL Data Library, April 1992
A;Description: Characterization of the murine macrophage mannose receptor: Demonstration
on.
A;Reference number: S21320
A;Accession: S21320
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-302,'W',303-1117,'E',1119-1455 <HA2>
A;Cross-references: EMBL:Z11974; NID:952997; PIDN:CAA78028.1; PID:G52998
R;Harris, N.; Peters, L.L.; Eicher, E.M.; Rits, M.; Raspberry, D.; Eichbaum, Q.G.; Super
Biochem. Biophys. Res. Commun. 198, 682-692, 1994
A;Title: The exon-intron structure and chromosomal localization of the mouse macrophage
A;Reference number: PC2245; MUID:94128116; PMID:8297379
A;Accession: PC2245
A;Molecule type: mRNA
A;Residues: 35-105 <HA3>
C;Genetics:
A;Gene: Mrcl
A;Map position: 2
C;Superfamily: phospholipase A2 receptor; C-type lectin homology; fibronectin type II re
C;Keywords: membrane protein; receptor
F;168-209/Domain: fibronectin type II repeat homology <2F9>
F;361-485/Domain: C-type lectin homology <LCH1>
F;943-1077/Domain: C-type lectin homology <LCH2>

Query Match 16.2%; Score 178.5; DB 1; Length 1455;
Best Local Similarity 25.9%; Pred. No. 1.7e-07;
Matches 53; Conservative 33; Mismatches 66; Indels 53; Gaps 8;

QY 18 YKVIYFHTSRRLNFEAEKACRRDGGQLVSISEDEQKLIKFIENLLPSDGFWIGLR 77
Db 806 YKDYQYFSKETMDNARRFCKKNFGDLATIKSESEKFLWKYI-NKNGGQSPYFIGML 864

QY 78 RREEKQSNSTACQDLYAWTDGSIQFRNWWYVDEPSCGS--EVCVVMYHOPSAPAGIGGPY 135
Db 865 ISMDKK-----FIWMDGSKVDYFVAWATGEPNPFANDDENCVTMY-----TNSGF---- 907

QY 136 MFQWNDRCNMKNPFICK---YSDEKPAVPSRAEAGEETELTTPVLPEETQE----- 184
Db 908 ---WNDINCGYPNPFICQRHNSINATAMP-----TTPTPGCKEGWHLYKNK 953

QY 185 -----EDAKKTFKESREAAALNL 201
Db 954 CFKIFGFANEKKSWQDARQACKGL 978

RESULT 3
S52781
neurocan - mouse
C;Species: Mus musculus (house mouse)
C;Date: 19-May-1995 #sequence_revision 21-Jul-1995 #text_change 04-Feb-2000
C;Accession: S52781
R;Rauch, U.; Forsberg, N.; Kulbe, G.; Arnold-Ammer, I.; Faessler, R.
submitted to the EMBL Data Library, February 1995
A;Description: Amino acid sequence of mouse neurocan and brevican and their different e
A;Reference number: S52781
A;Accession: S52781
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1268 <RAU>
A;Cross-references: EMBL:X84727; NID:9758629; PIDN:CAA59216.1; PID:9758630
C;Superfamily: aggregan; C-type lectin homology; complement factor H repeat homology; E
F;176-253/Domain: link protein repeat homology <LNK1>
F;274-355/Domain: link protein repeat homology <LNK2>
F;964-995/Domain: EGF homology <EGF>
F;1040-1160/Domain: C-type lectin homology <LCH>
F;1167-1223/Domain: complement factor H repeat homology <FHD>

Query Match 16.1%; Score 177; DB 2; Length 1268;
Best Local Similarity 31.5%; Pred. No. 1.9e-07;
Matches 45; Conservative 17; Mismatches 49; Indels 32; Gaps 6;

QY 14 QRPCYKVIYFHTSRRLNFEAEKACRRDGGQLVSISEDEQKLIKFIENLLPSDGDWF 73
Db 1048 QGHCYR--YF---AHRRAWEDAERCRRRAGHLTSVHSPEEHKFINSF-----GHENSW 1096

QY 74 IGLRRREEKQSNSTACQDLYAWTDGSIQFRNWWYVDEPS---CGSEVCVVMYHOPSAPAG 130
Db 1097 IGLNDRTVRD-----FQWTDNTGLQYENWREKQPDNFFAGGEDCVVMVAHESG--- 1145

QY 131 IGGPYMFQWNDRCNMKNPFICK 153
Db 1146 -----RWNDVPCNLYNPYVCK 1161

RESULT 4
T14274
versican precursor, splice form V2 - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 05-May-2000
C;Accession: T14274
R;Schmalfeldt, M.; Dours-Zimmermann, M.T.; Winterhalter, K.H.; Zimmermann, D.R.
J. Biol. Chem. 273, 15758-15764, 1998
A;Title: Versican V2 is a major extracellular matrix component of the mature bovine bra
A;Reference number: Z17954; MUID:98288320; PMID:9624174
A;Accession: T14274
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1643 <SCH>
A;Cross-references: EMBL:AF060458; NID:93253303; PID:93253304; PIDN:AAC24360.1
A;Experimental source: brain
C;Keywords: glycoprotein
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-1643/Product: versican, splice form V2 #status predicted <MAT>
F;57,331,352,817,965,1017,1333,1616,1626/Binding site: carbohydrate (Asn) (covalent) #s

Query Match 15.9%; Score 174.5; DB 2; Length 1643;
Best Local Similarity 25.5%; Pred. No. 4.4e-07;
Matches 59; Conservative 26; Mismatches 71; Indels 75; Gaps 11;

QY 14 QRPCYKVIYFHTSRRLNFEAEKACRRDGGQLVSISEDEQKLIKFIENLLPSDGDGF- 72
Db 1424 QGCYK--YF---AHRRTWDAARECRLOGAHLTSILSHBEQMFNRV-----GHDYQ 1471

Db	1472	WIGL-----NDKMFEDFRWTDGSTLQYENWRPNQDPSFFSTGTGDCVVIWHENG--	1521
Qy	129	AGIGGPYMFQNDRRCNMKNFICKYS-----DEKPAVPSPREAGE-----	169
Db	1522	-----QWNDRVPCNYHLTYTCCKGTACGQPPWVENAKTFGKMKPRYEINSLIRYHC	1572
Qy	170	-----ETELTT-----PVL-----PRETQEEDA K T F K E S R E A A L N	200
Db	1573	KDGFIO R H L P T I R C L G N R W A M P K I T C L N P S A Y Q R T Y S K K Y F K N S S A K D N	1623

RESULT 5

T42389

versican precursor, splice form V0 - bovine
N;Alternate names: chondroitin sulfate proteoglycan
C;Species: Bos primigenius taurus (cattle)
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 05-May-2000
C;Accession: T42389
R;Schmalfeldt, M.; Dours-Zimmermann, M.T.; Winterhalter, K.H.; Zimmermann, D.R.
J. Biol. Chem. 273, 15758-15764, 1998
A;Title: Versican V2 is a major extracellular matrix component of the mature bovine brain
A;Reference number: Z17954; MUID:98288320; PMID:9624174
A;Accession: T42389

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-3381 <S

A;C:Cross-references: EMBL:AF060456; NID:g32533299; PID:g3253300; PIDN:AAC24358.1
C;Superfamily: chicken chondroitin sulfate proteoglycan PG-M core protein; C-type lectin
C;Keywords: chondroitin sulfate proteoglycan; extracellular matrix; glycoprotein
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-3381/Product: versican, splice form V0 #status predicted <MAT>

Query Match

Best Local Similarity 25.5%: Pred. No. 1e-06:

Matches 59; Conservative 26; Mismatches

[illegible]

٩٥

Db 3162 QGQCYK--YF---AHRRTWDAARECRLOGAHLTSILSHEEQMFNRV-----GHYQ 3209

```

QY      73 WIGLRRREEKQSNSTACQDLYAWTDGSIQFNWYVDEP---SCGSEVCVMYHQPSAP 128
      |||||      |      :      :      :      :      :      :      :      :

```

DB 3210 WIGL-----NDKMFHDFRWTGSLQYENWRPNQPDFSTGEDCVIIWHENG-- 3259

128 ACICCPYMEQWINDDDCMMWBIETCWC

129 AGIGGEIMFQWNUDRCNMKNFICKYS----DEKPAVPSREAGE----- 169

Db 3260 -----OWNDPVCNYHLTYTCKKGTVACGPPWVENAKTEGKMKRPVEINSLPYHC 3310

K M N E T U C R I T H Z A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

QY 170 -----ETELT-----PVL-----PEETQEEDAKTFKESREALN 200

2

DD 3311 KDGFIQRHLPTIRCLNGRWPMPKITCLNPSAYQRTYSKKYFKNSSAKDN 3361

RESULT 6

\$28764

neurocan precursor - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 22-Nov-1993 #sequence_revision 01-Sep-1995 #text change 04-Feb-2000

C;Accession: S28764

R; Rauch, U.; Karthikeyan, L.; Maurel, P.; Margolis, R.U.; Margolis, R.K.

J. Biol. Chem. 267, 19536-19547, 1992

A;Title: Cloning and primary structure of neurocan, a developmentally regulated, aggrega

A;Reference number: S28764; MUID:92406907; PMID:1326557

A;Accession: S28764

A; Molecule type: mRNA

A;Resiques: I-1257 <RAU>

A;CROSS-References: EMBL:M97161; NID:g205649; PIDN:AAC37679.1; PID:g205650

[illegible]

Key words: chondroitin sulfate proteoglycan; glycoprotein

| P:1-22/Domain. signal semionce #otatue nradinted -012-

RESULT 8
A60979
versican precursor - human
N;Alternate names: chondroitin sulfate proteoglycan 2; chondroitin sulfate proteoglycan
N;Contents: glial hyaluronate-binding protein
C;Species: Homo sapiens (man)
C;Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text_change 19-Jan-2001
C;Accession: S06014; S43921; A60979; A30358; A29348; A45131; I54179
R;Zimmermann, D.R.; Ruoslahti, E.
EMBO J. 8, 2975-2981, 1989
A;Title: Multiple domains of the large fibroblast proteoglycan, versican.
A;Reference number: S06014; MUID:90059882; PMID:2583089
A;Accession: S06014
A;Molecule type: mRNA
A;Residues: 1-2409 <ZIM>
A;Cross-references: GB:X15998; NID:g37662; PIDN:CAA34128.1; PID:g37663
R;Yao, L.Y.; Moody, C.; Schoenherr, E.; Wight, T.N.; Sandell, L.J.
Matrix Biol. 14, 213-225, 1994
A;Title: Identification of the proteoglycan versican in aorta and smooth muscle cells by
A;Reference number: S43921; MUID:95005762; PMID:7921538
A;Accession: S43921
A;Molecule type: mRNA
A;Residues: 208-440;1094-1385;1910-2246 <YAO>
R;Bignami, A.; Lane, W.S.; Andrews, D.; Dahl, D.
Brain Res. Bull. 22, 67-70, 1989
A;Title: Structural similarity of hyaluronate binding proteins in brain and cartilage.
A;Reference number: A60979; MUID:89229983; PMID:2469524
A;Accession: A60979
A;Molecule type: protein
A;Residues: 171-210;289-303 <BIG>
R;Perides, G.; Lane, W.S.; Andrews, D.; Dahl, D.; Bignami, A.
J. Biol. Chem. 264, 5981-5987, 1989
A;Title: Isolation and partial characterization of a glial hyaluronate-binding protein.
A;Reference number: A30358; MUID:89174663; PMID:2466833
A;Accession: A30358
A;Molecule type: protein
A;Residues: 24-50;80-87,'D',89-119;128-155;167-218;229-259,'IR',261-268;277-283,'G',285-
R;Krusius, T.; Gehlsen, K.R.; Ruoslahti, E.
J. Biol. Chem. 262, 13120-13125, 1987
A;Title: A fibroblast chondroitin sulfate proteoglycan core protein contains lectin-like
A;Reference number: A29348; MUID:88007514; PMID:2820964
A;Accession: A29348
A;Molecule type: mRNA
A;Residues: 1725,'V',1727-2409 <KRU>
A;Cross-references: GB:J02814
R;Perides, G.; Rahemtulla, F.; Lane, W.S.; Asher, R.A.; Bignami, A.
J. Biol. Chem. 267, 23883-23887, 1992
A;Title: Isolation of a large aggregating proteoglycan from human brain.
A;Reference number: A45131; MUID:93054750; PMID:1429726
A;Contents: brain
A;Accession: A45131
A;Molecule type: protein
A;Residues: 21-22,'X',24-37 <PE2>
A;Experimental source: brain
A;Note: sequence extracted from NCBI backbone (NCBIP:118884)
R;Iozzo, R.V.; Naso, M.F.; Cannizzaro, L.A.; Wasmuth, J.J.; McPherson, J.D.
Genomics 14, 845-851, 1992
A;Title: Mapping of the versican proteoglycan gene (CSPG2) to the long arm of human chro
A;Reference number: I54179; MUID:93122792; PMID:1478664
A;Accession: I54179
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 251-347 <RES>
A;Cross-references: GB:S52488; NID:g263313; PIDN:AAB24878.1; PID:g263314
C;Genetics:
A;Gene: GDB:CSPG2
A;Cross-references: GDB:127873; OMIM:118661
A;Map position: 5q12-5q14
C;Superfamily: versican; C-type lectin homology; complement factor H repeat homology; EG
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-2409/Product: proteoglycan 24K core protein #status predicted <MAT>
F;167-244/Domain: link protein repeat homology <LNK1>

F;265-346/Domain: link protein repeat homology <LNK2>
F;559-1654/Domain: chondroitin sulfate attachment #status predicted <GAG>
F;2106-2137/Domain: EGF homology <EG1>
F;2144-2175/Domain: EGF homology <EG2>
F;2182-2302/Domain: C-type lectin homology <LCH>
F;2309-2365/Domain: complement factor H repeat homology <FHD>
Query Match 15.8%; Score 174; DB 1; Length 2409;
Best Local Similarity 28.5%; Pred. No. 7.5e-07;
Matches 47; Conservative 23; Mismatches 55; Indels 40; Gaps 8;
QY 14 QRPCYKVIYFHDTSRRLLNFEEAKEACRRDGGQLVSI ESEDEQKLIKFIENLLPSDGDGDF- 72
| | | | | : : : : : | | | | | : : : : :
Db 2190 QGQCYK--YF---AHRRTWDAERECLQGAHLTSLSHEEQVFNRV-----GHDYQ 2237
QY 73 WIGLRRREKQSNSTACQDLYAWTDGSI SQFRNWWYVDEP-----SCGSEVCVVMYHQSAP 128
| | | | | : : : : : | | | | | : : : : :
Db 2238 WIGL-----NDKMFEDFRWTDGSTLQYENWRPNQPDSPFSAGEDCVWIWHENG-- 2287
QY 129 AGIGGPMFQWDDRCNMKNFICKYS---DEKPAVPSREAEGE 169
| | | | | : : : : : | | | | | : : : : :
Db 2288 -----QWNDVPCNVHLYTYTCKGTGTVACGQPPVVENAKTFGK 2323
RESULT 9
A47171
chondroitin sulfate proteoglycan PG-M core protein - chicken
C;Species: Gallus gallus (chicken)
C;Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 21-Jul-2000
C;Accession: A47171
R;Shinomura, T.; Nishida, Y.; Ito, K.; Kimata, K.
J. Biol. Chem. 268, 14461-14469, 1993
A;Title: CDNA cloning of PG-M, a large chondroitin sulfate proteoglycan expressed durin
A;Reference number: A47171; MUID:93300846; PMID:8314802
A;Accession: A47171
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-3562 <SHI>
A;Cross-references: GB:D13542; NID:g391643; PIDN:BAA02742.1; PID:g391644
A;Experimental source: stage 22-23 developing limb buds
A;Note: sequence extracted from NCBI backbone (NCBIN:134456, NCBIP:134457)
C;Superfamily: chicken chondroitin sulfate proteoglycan PG-M core protein; C-type lectin
F;166-243/Domain: link protein repeat homology <LNK1>
F;264-345/Domain: link protein repeat homology <LNK2>
F;3258-3289/Domain: EGF homology <EGF1>
F;3296-3327/Domain: EGF homology <EGF>
F;3334-3454/Domain: C-type lectin homology <LCH>
F;3461-3517/Domain: complement factor H repeat homology <FHD>
Query Match 15.5%; Score 171; DB 2; Length 3562;
Best Local Similarity 28.5%; Pred. No. 2.2e-06;
Matches 47; Conservative 23; Mismatches 55; Indels 40; Gaps 8;
QY 14 QRPCYKVIYFHDTSRRLLNFEEAKEACRRDGGQLVSI ESEDEQKLIKFIENLLPSDGDGDF- 72
| | | | | : : : : : | | | | | : : : : :
Db 3342 QGQCYK--YF---AHRRTWDTAERECLQGAHLTSLSHEEQVFNRI-----GHDYQ 3389
QY 73 WIGLRRREKQSNSTACQDLYAWTDGSI SQFRNWWYVDEP-----SCGSEVCVVMYHQSAP 128
| | | | | : : : : : | | | | | : : : : :
Db 3390 WIGL-----NDKMFEDFRWTDGSPLOQYENWRPNQPDSPFSAGEDCVWIWHENG-- 3439
QY 129 AGIGGPMFQWDDRCNMKNFICKYS---DEKPAVPSREAEGE 169
| | | | | : : : : : | | | | | : : : : :
Db 3440 -----QWNDVPCNVHLYTYTCKGTGTVACGQPPVVENAKTFGK 3475
RESULT 10
T42710
mannose receptor, macrophage - mouse
N;Alternate names: lambda lectin; phospholipase A2 receptor
C;Species: Mus musculus (house mouse)
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jun-2000

C;Accession: T42710
R;Wu, K.; Yuan, J.; Lasky, L.A.
J. Biol. Chem. 271, 21323-21330, 1996
A;Title: Characterization of a novel member of the macrophage mannose receptor type C 1e
A;Reference number: 222235; MUID:96355501; PMID:8702911
A;Accession: T42710
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1479 <WUK>
A;Cross-references: EMBL:U56734; NID:G1336073; PID:G1336074; PIDN:AAC52729.1
C;Superfamily: phospholipase A2 receptor; C-type lectin homology; fibronectin type II re
C;Keywords: membrane protein; receptor
F;186-227/Domain: fibronectin type II repeat homology <2PR>

Query Match 15.5%; Score 170.5; DB 2; Length 1479;
Best Local Similarity 31.4%; Pred. No. 8.7e-07;
Matches 48; Conservative 22; Mismatches 48; Indels 35; Gaps 7;

QY 4 SQPVCRCGGTQPCYKVIYFHDTSRRLLNFEEAKEACRRDGGQLVSIIESEDEQKLIIEKFI 63
Db 384 SWQPF-----QGHCYRL-----QAEKRSWQESKRACLRGGDLLSIHSMAELEFITKQIK 433

QY 64 NLLPSDGDFFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNWWYVDEPS---CGSEVCVV 120
Db 434 QEVE---ELWIGL-----NDLKLQMNFEWSGSLVSFTHWHFFEPNFRDSLEDVCT 482

QY 121 MYHQPSAPAGIGGPPYMFQWDDRCNMKNPFICK 153
Db 483 IW-----GPEG-----RWNDSPCNQSLPSICK 504

RESULT 11

A39808
proteoglycan core protein, cartilage - bovine (fragments)
N;Alternate names: aggrecan; aggregating cartilage proteoglycan
C;Species: Bos primigenius taurus (cattle)
C;Date: 20-Mar-1992 #sequence revision 23-Mar-1995 #text change 13-Aug-1999
C;Accession: A34234; A27752; A39808; A27751; E29164; B27751; C27751; D27751; E27751; F27
R;Antonsson, P.; Heinigard, D.; Oldberg, A.
J. Biol. Chem. 264, 16170-16173, 1989
A;Title: The keratan sulfate-enriched region of bovine cartilage proteoglycan consists c
A;Reference number: A34234; MUID:89380219; PMID:2528543
A;Accession: A34234
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 128-621 <ANT>
A;Cross-references: GB:J05028
R;Oldberg, A.; Antonsson, P.; Heinigard, D.
Biochem. J. 243, 255-259, 1987
A;Title: The partial amino acid sequence of bovine cartilage proteoglycan, deduced from a
A;Reference number: A27752; MUID:87270630; PMID:3111460
A;Accession: A27752
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 622-1340 <OLD>
R;Sandy, J.D.; Boynton, R.E.; Flannery, C.R.
J. Biol. Chem. 266, 8198-8205, 1991
A;Title: Analysis of the catabolism of aggrecan in cartilage explants by quantitation of
A;Reference number: A39808; MUID:91217051; PMID:2022637
A;Accession: A39808
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-28;59-82;131-137,'QSET',142-149;196-207;226-249;1137-1143;1252-1267;1274-1
R;Perin, J.P.; Bonnet, F.; Jolles, P.
FEBS Lett. 206, 73-77, 1986
A;Title: Structural relationship between link proteins and proteoglycan monomers.
A;Reference number: A27751; MUID:87005253; PMID:3530809
A;Accession: A27751
A;Molecule type: protein
A;Residues: 29-58;74-130;174-175,'A',177-204;208-225 <PER>
R;Perin, J.P.; Bonnet, F.; Jolles, P.
FEBS Lett. 176, 37-42, 1984
A;Title: Sequence data concerning the protein core of the cartilage proteoglycan monomer

A;Reference number: A91327; MUID:85027710; PMID:6489519
A;Accession: E29164
A;Molecule type: protein
A;Residues: 1230-1249 <PE2>
C;Superfamily: aggrecan; C-type lectin homology; complement factor H repeat homology;
C;Keywords: glycoprotein
F;8-28/Domain: link protein repeat homology (fragment) <LNK1>
F;29-58/Domain: link protein repeat homology (fragment) <LNK2>
F;80-146/Domain: link protein repeat homology (fragments) <LNK3>
F;167-248/Domain: link protein repeat homology <LNK4>
F;1130-1250/Domain: C-type lectin homology <LCH>
F;1257-1313/Domain: complement factor H repeat homology <FHD>

Query Match 14.4%; Score 158.5; DB 2; Length 1340;
Best Local Similarity 26.8%; Pred. No. 9.1e-06;
Matches 48; Conservative 29; Mismatches 59; Indels 43; Gaps 10;

QY 6 QPVCRCGG---TORPCYKVIYFHDTSRRLLNFEEAKEACRRDGGQLVSIIESEDEQKLIIEKFI 62
Db 1127 QKLCEEGWTKFQGHCHYR--HFPD---RATWVDAESQCRKQKQSHLSSIVTPPEQ----EFV 1177

QY 63 ENLLPSDGDGF-WIGLRRREEKQSNSTACQDLYAWTDGSIQFRNWWYVDEP---SCGSEV 117
Db 1178 NN---NAQDYQWIGL-----NDKTIEGDFRWSDGHSLOFENWRPNQDNFFATGEDC 1226

QY 118 CVVMYHQPSAPAGIGGPPYMFQWDDRCNMKNPFICKYS-----DEKPAVPSREAEGETE 172
Db 1227 VVMIWHEKG-----EWNDVPCNYQLPFTCKKGTACGEPVVEHARIFGQKKD 1274

RESULT 12

T42630
aggrecan - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000
C;Accession: T42630
R;Hering, T.M.; Kollar, J.; Huynh, T.D.
submitted to the EMBL Data Library, September 1996
A;Description: Complete coding sequence of bovine aggrecan: comparative structural an
A;Reference number: 222182
A;Accession: T42630
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-2327 <HER>
A;Cross-references: EMBL:U76615; NID:G1730259; PID:G1730260; PIDN:AAB38524.1
A;Experimental source: articular chondrocytes
C;Superfamily: aggrecan; C-type lectin homology; complement factor H repeat homology;
C;Keywords: cartilage; chondroitin sulfate proteoglycan; extracellular matrix; glycop

Query Match 14.4%; Score 158.5; DB 2; Length 2327;
Best Local Similarity 26.8%; Pred. No. 1.7e-05;
Matches 48; Conservative 29; Mismatches 59; Indels 43; Gaps 10;

QY 6 QPVCRCGG---TORPCYKVIYFHDTSRRLLNFEEAKEACRRDGGQLVSIIESEDEQKLIIEKFI 62
Db 2114 QKLCEEGWTKFQGHCHYR--HFPD---RATWVDAESQCRKQKQSHLSSIVTPPEQ----EFV 2164

QY 63 ENLLPSDGDGF-WIGLRRREEKQSNSTACQDLYAWTDGSIQFRNWWYVDEP---SCGSEV 117
Db 2165 NN---NAQDYQWIGL-----NDKTIEGDFRWSDGHSLOFENWRPNQDNFFATGEDC 2213

QY 118 CVVMYHQPSAPAGIGGPPYMFQWDDRCNMKNPFICKYS-----DEKPAVPSREAEGETE 172
Db 2214 VVMIWHEKG-----EWNDVPCNYQLPFTCKKGTACGEPVVEHARIFGQKKD 2261

RESULT 13

A39086
aggrecan precursor, cartilage long splice form [validated] - human
N;Alternate names: chondroitin sulfate proteoglycan 1; large aggregating proteoglycan;
N;Contains: aggrecan cartilage short splice form
C;Species: Homo sapiens (man)
C;Date: 10-Sep-1999 #sequence_revision 01-Dec-2000 #text_change 08-Dec-2000

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 22, 2003, 16:07:10 ; Search time 19.3982 Seconds
(without alignments)
1954.412 Million cell updates/sec

Title: US-09-887-855-2_COPY_25_227
Perfect score: 1100
Sequence: 1 RLLSGQPVCRGGTQPCYKV.....EEDAKKTFKESREALNLAY 203

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 696363 seqs, 186758610 residues

Total number of hits satisfying chosen parameters: 696363

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA.*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1100	100.0	206	9	US-09-887-855-5
2	1100	100.0	374	9	US-09-887-855-2
3	1100	100.0	374	12	US-10-094-749-2090
4	1100	100.0	374	15	US-10-149-819-15
5	1096	99.6	374	12	US-10-094-749-2142
6	1086	98.7	382	10	US-09-909-320-137
7	1086	98.7	382	10	US-09-909-088B-137
8	1086	98.7	382	10	US-09-905-291A-137
9	1086	98.7	382	10	US-09-902-853-137
10	1086	98.7	382	10	US-09-907-824-137
11	1086	98.7	382	10	US-09-907-841-137
12	1086	98.7	382	11	US-09-904-011-137
13	1086	98.7	382	11	US-09-906-742-137
14	1086	98.7	382	11	US-09-906-838-137
15	1086	98.7	382	11	US-09-907-613-137

16	1086	98.7	382	11	US-09-907-942-137	Sequence 137, App
17	1086	98.7	382	11	US-09-904-859-137	Sequence 137, App
18	1086	98.7	382	11	US-09-909-204-137	Sequence 137, App
19	1086	98.7	382	11	US-09-904-820-137	Sequence 137, App
20	1086	98.7	382	11	US-09-904-786-137	Sequence 137, App
21	1086	98.7	382	11	US-09-906-646-137	Sequence 137, App
22	1086	98.7	382	11	US-09-906-700-137	Sequence 137, App
23	1086	98.7	382	11	US-09-903-786-137	Sequence 137, App
24	1086	98.7	382	11	US-09-902-903-137	Sequence 137, App
25	1086	98.7	382	11	US-09-903-749A-137	Sequence 137, App
26	1086	98.7	382	11	US-09-904-119-137	Sequence 137, App
27	1086	98.7	382	11	US-09-904-956-137	Sequence 137, App
28	1086	98.7	382	11	US-09-902-736-137	Sequence 137, App
29	1086	98.7	382	11	US-09-907-794-137	Sequence 137, App
30	1086	98.7	382	11	US-09-903-943-137	Sequence 137, App
31	1086	98.7	382	11	US-09-904-462-137	Sequence 137, App
32	1086	98.7	382	11	US-09-907-925-137	Sequence 137, App
33	1086	98.7	382	11	US-09-902-692-137	Sequence 137, App
34	1086	98.7	382	11	US-09-903-520-137	Sequence 137, App
35	1086	98.7	382	11	US-09-905-056-137	Sequence 137, App
36	1086	98.7	382	11	US-09-909-064-137	Sequence 137, App
37	1086	98.7	382	11	US-09-904-553-137	Sequence 137, App
38	1086	98.7	382	11	US-09-905-381-137	Sequence 137, App
39	1086	98.7	382	11	US-09-905-088-137	Sequence 137, App
40	1086	98.7	382	11	US-09-907-575-137	Sequence 137, App
41	1086	98.7	382	11	US-09-905-075-137	Sequence 137, App
42	1086	98.7	382	11	US-09-902-759-137	Sequence 137, App
43	1086	98.7	382	11	US-09-902-634-137	Sequence 137, App
44	1086	98.7	382	11	US-09-902-713-137	Sequence 137, App
45	1086	98.7	382	11	US-09-907-979-137	Sequence 137, App

ALIGNMENTS

RESULT 1
US-09-887-855-5
; Sequence 5, Application US/09887855
; Patent No. US20020058310A1
; GENERAL INFORMATION:
; APPLICANT: Immunex Corporation
; APPLICANT: Anderson, Dirk M
; TITLE OF INVENTION: LECTIN SS3939 DNA AND POLYPEPTIDES
; FILE REFERENCE: 2883-US
; CURRENT APPLICATION NUMBER: US/09/887,855
; CURRENT FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 206
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-887-855-5

Query Match	100.0%;	Score 1100;	DB 9;	Length 206;			
Best Local Similarity	100.0%;	Pred. No. 3.3e-103;					
Matches 203; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;			
Qy	1	RLLSGQPVCRGGTQPCYKVIFYHDT	SRRLNFEEAKEACRRDGGQLV	SI	SEDEQKLIK	60	
Db	4	RLLSGQPVCRGGTQPCYKVIFYHDT	SRRLNFEEAKEACRRDGGQLV	SI	SEDEQKLIK	63	
Qy	61	FIENLLPSDGD	FWIGLRRRREKQSN	TACQDLYAWTDG	SI	SQFRNMYVDEPSCGSEVCV	120
Db	64	FIENLLPSDGD	FWIGLRRRREKQSN	TACQDLYAWTDG	SI	SQFRNMYVDEPSCGSEVCV	123
Qy	121	MYHOPSAPAGIGG	PFQWMDRRCNMKN	FNICKYSDEKPAVPS	REA	GEETELTTPVLPE	180
Db	124	MYHOPSAPAGIGG	PFQWMDRRCNMKN	FNICKYSDEKPAVPS	REA	GEETELTTPVLPE	183
Qy	181	ETQEDAKKTFKES	REALNLAY	203			
Db	184	ETQEDAKKTFKES	REALNLAY	206			

```
RESULT 2
US-09-887-855-2
; Sequence 2, Application US/09887855
; Patent No. US20020058310A1
; GENERAL INFORMATION:
; APPLICANT: Immunex Corporation
; APPLICANT: Anderson, Dirk M
; TITLE OF INVENTION: LECTIN SS3939 DNA AND POLYPEPTIDES
; FILE REFERENCE: 2883-US
; CURRENT APPLICATION NUMBER: US/09/887,855
; CURRENT FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-887-855-2

Query Match      100.0%; Score 1100; DB 9; Length 374;
Best Local Similarity 100.0%; Pred. No. 7.3e-103;
Matches 203; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 RLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI ESEDEQKLI EK 60
Db      25 RLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI ESEDEQKLI EK 84

QY      61 FIENLLPSDGFWIGLRRRREKQSNSTACQDLYAWTDGSI SQFRNWWYVDEPSCGSEVCV 120
Db      85 FIENLLPSDGFWIGLRRRREKQSNSTACQDLYAWTDGSI SQFRNWWYVDEPSCGSEVCV 144

QY      121 MYHQSAPAGIGGPPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGETELTTPVLPE 180
Db      145 MYHQSAPAGIGGPPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGETELTTPVLPE 144

QY      181 ETQEEDAKKTFKESREAAALNLAY 203
Db      205 ETQEEDAKKTFKESREAAALNLAY 227
```

```
RESULT 3
US-10-094-749-2090
; Sequence 2090, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHIKO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOUYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH cDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
```

```
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2090
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-094-749-2090

Query Match      100.0%; Score 1100; DB 12; Length 374;
Best Local Similarity 100.0%; Pred. No. 7.3e-103;
Matches 203; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 RLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI ESEDEQKLI EK 60
Db      25 RLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI ESEDEQKLI EK 84

QY      61 FIENLLPSDGFWIGLRRRREKQSNSTACQDLYAWTDGSI SQFRNWWYVDEPSCGSEVCV 120
Db      85 FIENLLPSDGFWIGLRRRREKQSNSTACQDLYAWTDGSI SQFRNWWYVDEPSCGSEVCV 144

QY      121 MYHQSAPAGIGGPPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGETELTTPVLPE 180
Db      145 MYHQSAPAGIGGPPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGETELTTPVLPE 204

QY      181 ETQEEDAKKTFKESREAAALNLAY 203
Db      205 ETQEEDAKKTFKESREAAALNLAY 227
```

```
RESULT 4
US-10-149-819-15
; Sequence 15, Application US/10149819
; Publication No. US20030044913A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: YUE, Henry
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: TANG, Y. Tom
; APPLICANT: PATTERSON, Chandra
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: SHAH, Purvi
; APPLICANT: LAL, Preeti
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: BURFORD, Neil
; TITLE OF INVENTION: EXTRACELLULAR MATRIX AND CELL ADHESION MOLECULES
; FILE REFERENCE: PF-0760 PCT
; CURRENT APPLICATION NUMBER: US/10/149,819
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/172,852; 60/172,354
; PRIOR FILING DATE: 1999-12-10; 1999-12-16
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PERL Program
; SEQ ID NO 15
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030044913A1 3143411CD1
US-10-149-819-15
```

```
Query Match      100.0%; Score 1100; DB 15; Length 374;
Best Local Similarity 100.0%; Pred. No. 7.3e-103;
Matches 203; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 RLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI ESEDEQKLI EK 60
Db      25 RLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI ESEDEQKLI EK 84

QY      61 FIENLLPSDGFWIGLRRRREKQSNSTACQDLYAWTDGSI SQFRNWWYVDEPSCGSEVCV 120
Db      85 FIENLLPSDGFWIGLRRRREKQSNSTACQDLYAWTDGSI SQFRNWWYVDEPSCGSEVCV 144
```

Qy 121 MYHQSAPAGIGPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGEETELTPVLPE 180
Db 145 MYHQSAPAGIGPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGEETELTPVLPE 204

Qy 181 ETQEDAKTKTFKESREAAALNLAY 203
Db 205 ETQEDAKTKTFKESREAAALNLAY 227

RESULT 5
US-10-094-749-2142
; Sequence 2142, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHIKO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2142
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-094-749-2142

Query Match 99.6%; Score 1096; DB 12; Length 374;
Best Local Similarity 99.5%; Pred. No. 1.8e-102;
Matches 202; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RLLSGQPCVCGGTQPCYKVIYFHDTSRRLNFEAEACRRDGGQLVSI ESEDEQKLIK 60
Db 25 RLLSGQPCVCGGTQPCYKVIYFHDTSRRLNFEAEACRRDGGQLVSI ESEDEQKLIK 84

Qy 61 FIENLLPSDGFWIGLRRREKQSNSTACQDLYAWTDGSI SQFRNWWYVDEPSCGSEVCW 120
Db 85 FIENLLPSDGFWIGLRRREKQSNSTACQDLYAWTDGSI SQFRNWWYVDEPSCGSEVCW 144

Qy 121 MYHQSAPAGIGPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGEETELTPVLPE 180
Db 145 MYHQSAPAGIGPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGEETELTPVLPE 204

Qy 181 ETQEDAKTKTFKESREAAALNLAY 203
Db 205 ETQEDTKTKTFKESREAAALNLAY 227

RESULT 6
US-09-909-320-137
; Sequence 137, Application US/09909320

; Patent No. US20020132240A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/909,320
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-909-320-137

Query Match 98.7%; Score 1086; DB 10; Length 382;
Best Local Similarity 96.2%; Pred. No. 2e-101;
Matches 203; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 RLLS-----GQVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
Db 25 RLLSASDLRLRGQPVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 84

QY 53 DEQKLIKFIENLLPSDGDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNMYVDEPS 112
Db 85 DEQKLIKFIENLLPSDGDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNMYVDEPS 144

QY 113 CGSEVCVMYHQPSAPAGIGGYPYMFQWNDRCNMKNFICKYSDEKPAVPSREAEGETE 172
Db 145 CGSEVCVMYHQPSAPAGIGGYPYMFQWNDRCNMKNFICKYSDEKPAVPSREAEGETE 204

QY 173 LTTPLPEETQEEDAKKTFKESREAAALNLAY 203
Db 205 LTTPLPEETQEEDAKKTFKESREAAALNLAY 235

RESULT 7
US-09-909-088B-137
; Sequence 137, Application US/09909088B
; Patent No. US20020146709A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/909,088B
; CURRENT FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089

; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-909-088B-137

Query Match 98.7%; Score 1086; DB 10; Length 382;
Best Local Similarity 96.2%; Pred. No. 2e-101;
Matches 203; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 RLLS-----GQVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
Db 25 RLLSASDLRLRGQPVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 84

QY 53 DEQKLIKFIENLLPSDGDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNMYVDEPS 112
Db 85 DEQKLIKFIENLLPSDGDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNMYVDEPS 144

QY 113 CGSEVCVMYHQPSAPAGIGGYPYMFQWNDRCNMKNFICKYSDEKPAVPSREAEGETE 172
Db 145 CGSEVCVMYHQPSAPAGIGGYPYMFQWNDRCNMKNFICKYSDEKPAVPSREAEGETE 204

QY 173 LTTPLPEETQEEDAKKTFKESREAAALNLAY 203
Db 205 LTTPLPEETQEEDAKKTFKESREAAALNLAY 235

RESULT 8
US-09-905-291A-137
; Sequence 137, Application US/09905291A
; Patent No. US20020160374A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel

APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William, I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: 10466-14
CURRENT APPLICATION NUMBER: US/09/905,291A
CURRENT FILING DATE: 2001-07-12
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: US 60/143,048
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: US 60/145,698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: US 60/146,222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/20944
PRIOR FILING DATE: 1999-09-13
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21547
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/23089
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: PCT/US99/28214
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: PCT/US99/28564
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/28565
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: PCT/US99/30911
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US99/30999
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US00/00219
NUMBER OF SEQ ID NOS: 423
SEQ ID NO 137
LENGTH: 382
TYPE: PRT
ORGANISM: Homo sapiens
US-09-905-291A-137

Query Match 98.7%; Score 1086; DB 10; Length 382;
Best Local Similarity 96.2%; Pred. No. 2e-101;
Matches 203; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
QY 1 RLLS-----GQVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESE 52
Db |||||
25 RLLSASDLRGCGQPCVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIESE 84
QY 53 DEQKLEKFIENLLPSDGFWIGLRRREKQSNSTACQDLYAWTDGSIQFRNWWYVDEPS 112
Db |||||
85 DEQKLEKFIENLLPSDGFWIGLRRREKQSNSTACQDLYAWTDGSIQFRNWWYVDEPS 144
QY 113 CGSEVCVMYHQPSAPAGIGGPMFQWDDRCNMKNFICKYSDEKPAVPSREAGEETE 172
Db |||||
145 CGSEVCVMYHQPSAPAGIGGPMFQWDDRCNMKNFICKYSDEKPAVPSREAGEETE 204
QY 173 LTTPLVPEETQEDAKKTFKESREAAINLAY 203
Db |||||
205 LTTPLVPEETQEDAKKTFKESREAAINLAY 235

RESULT 9
US-09-902-853-137
Sequence 137, Application US/09902853
Publication No. US20020192659A1

GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
APPLICANT: Ashkenazi, Avi
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, A.
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth, J.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Mather, Jennie P.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William, I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: 10466-14
CURRENT APPLICATION NUMBER: US/09/902,853
CURRENT FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: US/09/665,350
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: US 60/143,048
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: US 60/145,698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: US 60/146,222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/20944
PRIOR FILING DATE: 1999-09-13
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21547
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/23089
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: PCT/US99/28214
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: PCT/US99/28564
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/28565
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: PCT/US99/30911
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US99/30999
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US00/00219
NUMBER OF SEQ ID NOS: 423
SEQ ID NO 137
LENGTH: 382
TYPE: PRT
ORGANISM: Homo Sapien
US-09-902-853-137

Query Match 98.7%; Score 1086; DB 10; Length 382;

```
Best Local Similarity 96.2%; Pred. No. 2e-101;
Matches 203; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

Qy 1 RLLS-----GQVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIIE 52
    |||||
Db 25 RLLSASDLRLGGQPVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIIE 84

Qy 53 DEQKLIIEFIENLLPSDGDWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNWWYVDEPS 112
    |||||
Db 85 DEQKLIIEFIENLLPSDGDWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNWWYVDEPS 144

Qy 113 CGSEVCVVMYHQPAPAGIGGYPYMFQWDDRCNMKNFNICKYSDEKPAVPSREAEGEETE 172
    |||||
Db 145 CGSEVCVVMYHQPAPAGIGGYPYMFQWDDRCNMKNFNICKYSDEKPAVPSREAEGEETE 204

Qy 173 LTPVLPEETQEEADAKKTFKESREAAALNLAY 203
    |||||
Db 205 LTPVLPEETQEEADAKKTFKESREAAALNLAY 235
```

```
RESULT 10
US-09-907-824-137
; Sequence 137, Application US/09907824
; Publication No. US20020197671A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/907,824
; CURRENT FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: 09/665,350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
```

```

; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-907-824-137

Query Match 98.7%; Score 1086; DB 10; Length 382;
Best Local Similarity 96.2%; Pred. No. 2e-101;
Matches 203; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

Qy 1 RLLS-----GQVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIIE 52
    |||||
Db 25 RLLSASDLRLGGQPVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIIE 84

Qy 53 DEQKLIIEFIENLLPSDGDWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNWWYVDEPS 112
    |||||
Db 85 DEQKLIIEFIENLLPSDGDWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNWWYVDEPS 144

Qy 113 CGSEVCVVMYHQPAPAGIGGYPYMFQWDDRCNMKNFNICKYSDEKPAVPSREAEGEETE 172
    |||||
Db 145 CGSEVCVVMYHQPAPAGIGGYPYMFQWDDRCNMKNFNICKYSDEKPAVPSREAEGEETE 204

Qy 173 LTPVLPEETQEEADAKKTFKESREAAALNLAY 203
    |||||
Db 205 LTPVLPEETQEEADAKKTFKESREAAALNLAY 235
```

```
RESULT 11
US-09-907-841-137
; Sequence 137, Application US/09907841
; Publication No. US20020198366A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
```

```
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/907,841
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-907-841-137

Query Match      98.7%; Score 1086; DB 10; Length 382;
Best Local Similarity 96.2%; Pred. No. 2e-101;
Matches 203; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY      1  RLLS-----GQVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
Db      25  RLLSASDLRLRGQVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 84

QY      53  DEQKLEKFIENLLPSDGDGFWIGLRRRREEKQSNSTACQDLYAWTDGSIQFRN 112
Db      85  DEQKLEKFIENLLPSDGDGFWIGLRRRREEKQSNSTACQDLYAWTDGSIQFRN 144

QY      113 CGSEVCVMYHQPAPAGIGPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGETE 172
Db      145 CGSEVCVMYHQPAPAGIGPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGETE 204

QY      173 LTTPLPEETQEEDAKTKFKESREAAALNLAY 203
Db      205 LTTPLPEETQEEDAKTKFKESREAAALNLAY 235

RESULT 12
US-09-904-011-137
; Sequence 137, Application US/09904011
; Publication No. US20030003530A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
```

```
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/904,011
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 09/665,350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-904-011-137
```

```
Query Match      98.7%; Score 1086; DB 11; Length 382;
Best Local Similarity 96.2%; Pred. No. 2e-101;
Matches 203; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY      1  RLLS-----GQVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
Db      25  RLLSASDLRLRGQVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 84

QY      53  DEQKLEKFIENLLPSDGDGFWIGLRRRREEKQSNSTACQDLYAWTDGSIQFRN 112
Db      85  DEQKLEKFIENLLPSDGDGFWIGLRRRREEKQSNSTACQDLYAWTDGSIQFRN 144
```


Qy 113 CGSEVCVVMYHQSAPAGIGGPPYMFQWDDRCNMKNPFICKYSDEKPAVPSREAEGETE 172
|||||
Db 145 CGSEVCVVMYHQSAPAGIGGPPYMFQWDDRCNMKNPFICKYSDEKPAVPSREAEGETE 204

Qy 173 LTPVLPEETOEDAKKTFKESREAAALNLAY 203
|||||
Db 205 LTPVLPEETOEDAKKTFKESREAAALNLAY 235

RESULT 13

US-09-906-742-137
; Sequence 137, Application US/09906742
; Publication No. US20030023054A1
; GENERAL INFORMATION:

; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/906,742
; CURRENT FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 09/665,350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02

; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-906-742-137

Query Match 98.7%; Score 1086; DB 11; Length 382;
Best Local Similarity 96.2%; Pred. No. 2e-101;
Matches 203; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

Qy 1 RLIS-----GQVCRGGTQPCYKVIYFHDTSRLNFEEAKEACRRDGGQLVSIIESE 52
|||||
Db 25 RLISASDLDLRGGQPCVCRGGTQPCYKVIYFHDTSRLNFEEAKEACRRDGGQLVSIIESE 84

Qy 53 DEQKLIKPIENLLPSDGFWIGLRRREEKQSNSTACODLYAWTDGSIQFRNWTVDPS 112
|||||
Db 85 DEQKLIKPIENLLPSDGFWIGLRRREEKQSNSTACODLYAWTDGSIQFRNWTVDPS 144

Qy 113 CGSEVCVVMYHQSAPAGIGGPPYMFQWDDRCNMKNPFICKYSDEKPAVPSREAEGETE 172
|||||
Db 145 CGSEVCVVMYHQSAPAGIGGPPYMFQWDDRCNMKNPFICKYSDEKPAVPSREAEGETE 204

Qy 173 LTPVLPEETOEDAKKTFKESREAAALNLAY 203
|||||
Db 205 LTPVLPEETOEDAKKTFKESREAAALNLAY 235

RESULT 14

US-09-906-838-137
; Sequence 137, Application US/09906838
; Publication No. US20030027143A1
; GENERAL INFORMATION:

; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.

; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/906,838
; CURRENT FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 09/665,350
; PRIOR FILING DATE: 2000-09-18

; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-906-838-137

Query Match 98.7%; Score 1086; DB 11; Length 382;
Best Local Similarity 96.2%; Pred. No. 2e-101;
Matches 203; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 RLLS-----GQVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIIESE 52
Db 25 RLLSASDLRLGGQPCVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIIESE 84

QY 53 DEQKLEKFIENLLPSDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNRYVDEPS 112
Db 85 DEQKLEKFIENLLPSDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNRYVDEPS 144

QY 113 CGSEVCVMYHQPSAPAGIGGPFQWDDRCNMKNMFICKYSDEKPAVPSRAEGEETE 172
Db 145 CGSEVCVMYHQPSAPAGIGGPFQWDDRCNMKNMFICKYSDEKPAVPSRAEGEETE 204

QY 173 LTTPVLPEETQEDAKKTFKESREAAALNLAY 203
Db 205 LTTPVLPEETQEDAKKTFKESREAAALNLAY 235

RESULT 15
US-09-907-613-137
; Sequence 137, Application US/09907613
; Publication No. US20030027145A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone

; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/907,613
; PRIOR FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-907-613-137

Query Match 98.7%; Score 1086; DB 11; Length 382;
Best Local Similarity 96.2%; Pred. No. 2e-101;
Matches 203; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 RLLS-----GQVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIIESE 52
Db 25 RLLSASDLRLGGQPCVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIIESE 84

Qy	53	DEQKLI	EFIE	NLPS	DGDF	WIGL	RRRE	EKQSN	STAC	QDLY	AWTD	GSIS	QFRN	WYVD	EPS	112
Db	85	DEQKLI	EFIE	NLPS	DGDF	WIGL	RRRE	EKQSN	STAC	QDLY	AWTD	GSIS	QFRN	WYVD	EPS	144
Qy	113	CGSEV	CVVM	YTHQ	PSAP	AGIG	PGPY	MFQW	NDRC	NMKNN	FICK	YSDE	KPAV	PSRE	AEETE	172
Db	145	CGSEV	CVVM	YTHQ	PSAP	AGIG	PGPY	MFQW	NDRC	NMKNN	FICK	YSDE	KPAV	PSRE	AEETE	204
Qy	173	LTPVL	PEET	QEED	AKKT	FKES	REAA	LN	LAY							203
Db	205	LTPVL	PEET	QEED	AKKT	FKES	REAA	LN	LAY							235

Search completed: December 22, 2003, 16:16:48
Job time : 20.3982 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 22, 2003, 16:05:09 ; Search time 6.4979 Seconds
(without alignments)
1476.391 Million cell updates/sec

Title: US-09-887-855-2_COPY_24_227
Perfect score: 1106
Sequence: 1 GRLLSGQPVCRGGTQPCYK.....BEDAKKTFKESREALNLAY 204

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues
Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	574	51.9	273	1 CHOD_MOUSE	Q9cxm0 mus musculus
2	561.5	50.8	273	1 CHOD_HUMAN	Q9h9p2 homo sapien
3	182	16.5	1456	1 MANR_HUMAN	P22897 homo sapien
4	177	16.0	1268	1 PGCN_MOUSE	P55066 mus musculus
5	176	15.9	1321	1 PGCN_HUMAN	O14594 homo sapien
6	174.5	15.8	3381	1 PGCN_BOVIN	P81282 bos taurus
7	174	15.7	1257	1 PGCN_RAT	P55067 rattus norv
8	174	15.7	2738	1 PGCN_RAT	Q9erb4 rattus norv
9	174	15.7	3358	1 PGCN_MOUSE	Q62059 mus musculus
10	174	15.7	3396	1 PGCN_HUMAN	P13611 homo sapien
11	171	15.5	3562	1 PGCN_CHICK	Q90953 gallus gall
12	165	14.9	643	1 CD93_RAT	Q9et61 rattus norv
13	158.5	14.3	2364	1 PGCA_BOVIN	P13608 bos taurus
14	158.5	14.3	2415	1 PGCA_HUMAN	P16112 homo sapien
15	155.5	14.1	2333	1 PGCA_CANFA	Q28343 canis famil
16	154.5	14.0	612	1 LEM2_MOUSE	Q00690 mus musculus
17	153.5	13.9	644	1 CD93_MOUSE	O89103 mus musculus
18	153.5	13.9	2124	1 PGCA_RAT	P07897 rattus norv
19	152	13.7	912	1 PGCB_BOVIN	Q28062 bos taurus
20	151.5	13.7	652	1 CD93_HUMAN	Q9npy3 homo sapien
21	151	13.7	197	1 CLF1_HUMAN	O75596 homo sapien
22	149	13.5	321	1 FCE2_HUMAN	P06734 homo sapien
23	147.5	13.3	2132	1 PGCA_MOUSE	Q61282 mus musculus
24	147	13.3	883	1 PGCB_MOUSE	Q61361 mus musculus
25	146	13.2	2109	1 PGCA_CHICK	P07898 gallus gall
26	145	13.1	158	1 LECG_TRIST	Q9ygp1 trimeresuru
27	145	13.1	883	1 PGCB_RAT	P55068 rattus norv
28	144.5	13.1	173	1 LEC2_MEGRO	P17346 megabalanus
29	144.5	13.1	372	1 LEM1_RAT	P30836 rattus norv
30	143.5	13.0	372	1 LEM1_MOUSE	P18337 mus musculus
31	141.5	12.8	331	1 FCE2_MOUSE	P20693 mus musculus
32	141.5	12.8	549	1 LEM2_RAT	P98105 rattus norv
33	140.5	12.7	162	1 LEC3_MEGRO	P07439 megabalanus

34	140	12.7	370	1 LEM1_BOVIN	P98131 bos taurus
35	139	12.6	248	1 PSPA_HUMAN	P07714 homo sapien
36	139	12.6	283	1 LECA_SARPE	P05047 sarcophaga
37	138.5	12.5	152	1 IXA_TRIFL	P23806 trimeresuru
38	138.5	12.5	202	1 TETN_MOUSE	P43025 mus musculus
39	137.5	12.4	372	1 LEM1_MACMU	Q95198 macaca mula
40	137.5	12.4	372	1 LEM1_PAPHA	Q28768 papio hamad
41	136	12.3	175	1 LITH_BOVIN	P23132 bos taurus
42	135	12.2	165	1 LIT1_MOUSE	P43137 mus musculus
43	134.5	12.2	132	1 ACAL_ANSAN	P83300 anser anser
44	134.5	12.2	372	1 LEM1_PONPY	Q95235 pongo pygma
45	134.5	12.2	485	1 LEM2_BOVIN	P98107 bos taurus

ALIGNMENTS

RESULT 1
CHOD_MOUSE
ID CHOD_MOUSE STANDARD; PRT; 273 AA.
AC Q9CXM0; Q8VI31;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Chondrolectin precursor (Transmembrane protein MT75).
GN CHODL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RA Weng L., Smits P., Hubner R., Wouters J., Merregaert J.;
RT "Mt75, a low expressed c-type lectin gene involving in
RT chondrogenesis";
RL Submitted (OCT-2000) to the EMBL/GenBank/DBDJ databases.
RN [2]
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryonic head;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombarts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

CC EMBL; AF311699; AAL50354.1; -.

DR Pfam; PF00059; lectin c; 1.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C TYPE LECTIN 1; FALSE_NEG.
DR PROSITE; PS0041; C_TYPE_LECTIN2; 1.
KW Lectin; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 273 CHONDROLECTIN.
FT DOMAIN 22 216 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 217 237 POTENTIAL.
FT DOMAIN 238 273 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 35 179 C-TYPE LECTIN.
FT CARBOHYD 86 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 273 AA; 30431 MW; F4890AABF572A311 CRC64;

Query Match 50.8%; Score 561.5; DB 1; Length 273;
Best Local Similarity 60.1%; Pred. No. 5.3e-43;
Matches 101; Conservative 25; Mismatches 35; Indels 7; Gaps 3;

Qy 2 RLLSGQVCRGGTQPCYKVIYFHDTSRRLLNFEEAKEACRRDGGQLVSIIESEDEQKLIK 61
Db 23 RVVSGQKVCFAFDKHPCKYKMAVPHLSRVSFQEARLACESEGGVLLSLENAEQKLIES 82

Qy 62 FIENLLP-----SDGDFWIGLRREEKQSNSTACQDLVATWTDGSIQFRNWWYVDEPSCGS 116
Db 83 MLQNLTKPGTGISDGDGFWIGLRNGDGT-SGACPDLYQWSDGNSQYRNWYTDEPSCGS 141

Qy 117 EVCVMYHQPSAPAGIGGPMFQWDDRCNMKNPFICKYSDB-KPAVP 163
Db 142 EKCVVMYHQPTANPGLGPPYLYQWDDRCNMKNHYICKYPEINPTAP 189

RESULT 3

MANR_HUMAN STANDARD; PRT; 1456 AA.
AC P22897;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Macrophage mannose receptor precursor (MMR) (CD206 antigen).
GN MRC1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Placenta;
RX MEDLINE=90324192; PubMed=2373685;
RA Taylor M.E., Conary J.T., Lennartz M.R., Stahl P.D., Drickamer K.;
RT "Primary structure of the mannose receptor contains multiple motifs
RT resembling carbohydrate-recognition domains.";
RL J. Biol. Chem. 265:12156-12162(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93052405; PubMed=1294118;
RA Kim S.J., Ruiz N., Bezouska K., Drickamer K.;
RT "Organization of the gene encoding the human macrophage mannose
RT receptor (MRC1).";
RL Genomics 14:721-727(1992).
RN [3]
RP STUDIES ON THE BINDING OF INDIVIDUAL LECTIN DOMAINS.
RX MEDLINE=92112893; PubMed=1730714;
RA Taylor M.E., Bezouska K., Drickamer K.;
RT "Contribution to ligand binding by multiple carbohydrate-recognition
RT domains in the macrophage mannose receptor.";
RL J. Biol. Chem. 267:1719-1726(1992).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 642-788.
RX MEDLINE=20347275; PubMed=10779515;
RA Feinberg H., Park-Snyder S., Kolatkar A.R., Heise C.T., Taylor M.E.,
RA Weis W.I.;
RT "Structure of a C-type carbohydrate recognition domain from the
RT macrophage mannose receptor.";

RL J. Biol. Chem. 275:21539-21548(2000).
CC -|- FUNCTION: MEDIATES THE ENDOCYTOSIS OF GLYCOPROTEINS BY
CC MACROPHAGES, IN SEVERAL RECOGNITION AND UPTAKE PROCESSES.
CC -|- SUBCELLULAR LOCATION: Type I membrane protein.
CC -|- MISCELLANEOUS: CRDS 1-3 HAVE AT MOST VERY WEAK AFFINITY FOR
CC CARBOHYDRATE. CRD 4 SHOWS THE HIGHEST AFFINITY BINDING AND HAS
CC MULTISPECIFICITY FOR A VARIETY OF MONOSACCHARIDES. AT LEAST 3 CRDS
CC (4, 5, AND 7) ARE REQUIRED FOR HIGH AFFINITY BINDING AND
CC ENDOCYTOSIS OF MULTIVALENT GLYCOCONJUGATES.
CC -|- SIMILARITY: Contains 8 C-type lectin family domains.
CC -|- SIMILARITY: Contains 1 ricin B-type lectin domain.
CC -|- DATABASE: NAME=PROW; NOTE=PROW 2:85-89(2001);
CC WWW="http://www.ncbi.nlm.nih.gov/prow/guide/1644341535_g.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; J05550; AAA59868.1; -
DR EMBL; M93221; AAA60389.1; -
DR EMBL; M93192; AAA60389.1; JOINED.
DR EMBL; M93193; AAA60389.1; JOINED.
DR EMBL; M93194; AAA60389.1; JOINED.
DR EMBL; M93195; AAA60389.1; JOINED.
DR EMBL; M93196; AAA60389.1; JOINED.
DR EMBL; M93197; AAA60389.1; JOINED.
DR EMBL; M93198; AAA60389.1; JOINED.
DR EMBL; M93199; AAA60389.1; JOINED.
DR EMBL; M93200; AAA60389.1; JOINED.
DR EMBL; M93201; AAA60389.1; JOINED.
DR EMBL; M93202; AAA60389.1; JOINED.
DR EMBL; M93203; AAA60389.1; JOINED.
DR EMBL; M93204; AAA60389.1; JOINED.
DR EMBL; M93205; AAA60389.1; JOINED.
DR EMBL; M93206; AAA60389.1; JOINED.
DR EMBL; M93207; AAA60389.1; JOINED.
DR EMBL; M93208; AAA60389.1; JOINED.
DR EMBL; M93209; AAA60389.1; JOINED.
DR EMBL; M93210; AAA60389.1; JOINED.
DR EMBL; M93211; AAA60389.1; JOINED.
DR EMBL; M93212; AAA60389.1; JOINED.
DR EMBL; M93213; AAA60389.1; JOINED.
DR EMBL; M93214; AAA60389.1; JOINED.
DR EMBL; M93215; AAA60389.1; JOINED.
DR EMBL; M93216; AAA60389.1; JOINED.
DR EMBL; M93217; AAA60389.1; JOINED.
DR EMBL; M93218; AAA60389.1; JOINED.
DR EMBL; M93219; AAA60389.1; JOINED.
DR EMBL; M93220; AAA60389.1; JOINED.
DR PIR; A36563; A36563.
DR PDB; 1EGG; 30-AUG-00.
DR PDB; 1EGI; 30-AUG-00.
DR Genew; HGNC:7228; MRC1.
DR MIM; 153618; -
DR GO; GO:0005887; C:integral to plasma membrane; TAS.
DR GO; GO:0005537; F:mannose binding activity; TAS.
DR GO; GO:0004872; F:receptor activity; TAS.
DR GO; GO:0006898; P:receptor mediated endocytosis; TAS.
DR InterPro; IPR002353; AntifreezeII.
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR000772; Ricin_B_lectin.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00059; lectin c; 8.
DR Pfam; PF00652; Ricin_B_lectin; 2.
DR PRINTS; PR00013; FNTYPEII.
DR PRINTS; PR00356; ANTIFREEZEII.
DR ProDom; PD000995; FN_Type_II; 1.
DR SMART; SM00034; CLECT; 8.


```
DR SMART; SM00032; CCP; 1.
DR SMART; SM00034; CLECT; 1.
DR SMART; SM00179; EGF_CA; 1.
DR SMART; SM00409; IG_1.
DR SMART; SM00445; LINK; 2.
DR PROSITE; PS00010; ASX_HYDROXYL; 1.
DR PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR PROSITE; PS00041; C_TYPE_LECTIN_2; 1.
DR PROSITE; PS00022; EGF_1; 3.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01187; EGF_CA; 1.
DR PROSITE; PS00835; IG_LINK; 1.
DR PROSITE; PS01241; LINK; 2.
KW Glycoprotein; Hyaluronic acid; Proteoglycan; Immunoglobulin domain;
KW EGF-like domain; Calcium; Repeat; Lectin; Sushi; Signal.
FT SIGNAL 1 22
FT CHAIN 23 1268
FT DOMAIN 37 157
FT DOMAIN 158 253
FT DOMAIN 259 355
FT DOMAIN 260 996
FT DOMAIN 998 1034
FT DOMAIN 1036 1165
FT DOMAIN 1166 1224
FT DISULFID 58 139
FT DISULFID 181 252
FT DISULFID 205 226
FT DISULFID 279 354
FT DISULFID 303 324
FT DISULFID 964 975
FT DISULFID 969 984
FT DISULFID 986 995
FT DISULFID 1040 1051
FT DISULFID 1068 1160
FT DISULFID 1136 1152
FT DISULFID 1167 1210
FT DISULFID 1196 1223
FT CARBOHYD 121 121
FT CARBOHYD 339 339
FT CARBOHYD 742 742
FT CARBOHYD 978 978
FT CARBOHYD 1175 1175
FT SEQUENCE 1268 AA; 137200 MW; 3014E8E202A2FAEC CRC64;

Query Match 16.0%; Score 177; DB 1; Length 1268;
Best Local Similarity 31.5%; Pred. No. 7.9e-08;
Matches 45; Conservative 17; Mismatches 49; Indels 32; Gaps 6;

QY 15 QRPCYKVIYFHTSRRLNFEAEKACRRDGGQLVSISEDEQKLEKFIENLLPSDGF 74
Db 1048 QGHCVYR--YF--AHRRWEDAEEDCRRRAGHLTVSHSPHEKFINSF-----GHENSW 1096
QY 75 IGLRRREEKQSNSTACQDLAWTDGSIQFRNWWYVDEPS---CGSEVCVVMYHQPSPAG 131
Db 1097 IGLNDRTVRD-----FQWTDNTGLQYENWREKQPDNFFAGGEDCVVMVAHESG--- 1145
QY 132 IGGPYMFQWDDRCNMKNFICK 154
Db 1146 -----RWNDVPCNPLPVCK 1161
```

RESULT 5

```
PGCN_HUMAN
ID PGCN_HUMAN STANDARD; PRT; 1321 AA.
AC O14594; Q9UPK6;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Neurocan core protein precursor (Chondroitin sulfate proteoglycan 3).
GN CSPG3 OR NCAN OR NEUR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

```
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99013874; PubMed=9795216;
RA Prange C.K., Pennacchio L.A., Lieuallen K., Fan W., Lennon G.G.;
RT "Characterization of the human neurocan gene, CSPG3.";
RL Gene 221:199-205(1998).
RN [2]
RP SEQUENCE OF 1-990 AND 1007-1321 FROM N.A.
RA Lamerdin J.E., McCready P.M., Skowronski E., Adamson A.W.,
RA Burkhardt-Schultz K., Gordon L., Kyle A., Ramirez M., Stilwagen S.,
RA Phan H., Velasco N., Do L., Regala W., Terry A., Ganes J.,
RA Danganan L., Poundstone P., Christensen M., Georgescu A., Avila J.,
RA Liu S., Attix C., Andreise T., Frankheim M., Amico-Keller G.,
RA Coefield J., Duarte S., Lucas S., Bruce R., Thomas P., Quan G.,
RA Kronmiller B., Arellano A., Montgomery M., Ow D., Nolan M., Trong S.,
RA Kobayashi A., Olsen A.S., Carrano A.V.;
RT "Sequence analysis of an ~1 Mb region containing the MEF2B gene in
19p12."
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: May modulate neuronal adhesion and neurite growth during
development by binding to neural cell adhesion molecules (NG-CAM
and N-CAM). Chondroitin sulfate proteoglycan; binds to hyaluronic
acid.
CC -!- TISSUE SPECIFICITY: Brain.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -!- SIMILARITY: Contains 2 EGF-like domains.
CC -!- SIMILARITY: Contains 2 link domains.
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC -!- SIMILARITY: Contains 1 Sushi (SCR) domain.
CC -!- SIMILARITY: BELONGS TO THE AGGRECAN/VERSICAN PROTEOGLYCAN FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; AF026547; AAC80576.1; -.
EMBL; AC003110; AAB86655.1; -.
EMBL; AC005254; AAC25581.1; -.
HSSP; P00740; 1EDM.
Genew; HGNC:2465; CSPG3.
MIM; 600826; -.
InterPro; IPR000152; Asx_hydroxyl.
InterPro; IPR000742; EGF_2.
InterPro; IPR001881; EGF_Ca.
InterPro; IPR006209; EGF_like.
InterPro; IPR007110; Ig-Like.
InterPro; IPR003599; Ig.
InterPro; IPR003006; Ig_MHC.
InterPro; IPR001304; Lectin_C.
InterPro; IPR000538; Link.
InterPro; IPR000436; Sushi_SCR_CCP.
Pfam; PF00008; EGF; 2.
Pfam; PF00047; ig; 1.
Pfam; PF00059; lectin_c; 1.
Pfam; PF00084; sushi; 1.
Pfam; PF00193; Xlink; 2.
PRINTS; PR01265; LINKMODULE.
ProDom; PD000918; Link; 2.
SMART; SM00032; CCP; 1.
SMART; SM00034; CLECT; 1.
SMART; SM00179; EGF_CA; 1.
SMART; SM00409; IG; 1.
SMART; SM00445; LINK; 2.
PROSITE; PS00010; ASX_HYDROXYL; 1.
PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
PROSITE; PS00041; C_TYPE_LECTIN_2; 1.
PROSITE; PS00022; EGF_1; 3.
PROSITE; PS01186; EGF_2; 1.
```


DR Pfam; PF00193; Xlink; 2.
DR PRINTS; PR01265; LINKMODULE.
DR PRINTS; PR00356; ANTIFREEZEII.
DR ProDom; PD000918; Link; 2.
DR SMART; SM00032; CCP; 1.
DR SMART; SM00034; CLECT; 1.
DR SMART; SM00179; EGF_CA; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00445; LINK; 2.
DR PROSITE; PS00010; ASX HYDROXYL; 1.
DR PROSITE; PS00615; C-TYPE_LECTIN_1; 1.
DR PROSITE; PS50041; C-TYPE_LECTIN_2; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01187; EGF_CA; 1.
DR PROSITE; PS50835; IG LIKE; 1.
DR PROSITE; PS01241; LINK; 2.
KW Glycoprotein; Proteoglycan; Lectin; Extracellular matrix; Sushi;
KW Signal; Repeat; EGF-like domain; Calcium; Immunoglobulin domain;
KW Hyaluronic acid; Alternative splicing.
FT SIGNAL 1 20
FT CHAIN 21 3381
FT DOMAIN 21 147
FT DOMAIN 168 245
FT DOMAIN 266 347
FT DOMAIN 349 1336
FT DOMAIN 1337 3074
FT DOMAIN 3074 3110
FT DOMAIN 3112 3148
FT DOMAIN 3161 3275
FT DOMAIN 3280 3338
FT DISULFID 44 131
FT DISULFID 173 244
FT DISULFID 197 218
FT DISULFID 271 346
FT DISULFID 295 316
FT DISULFID 3078 3089
FT DISULFID 3083 3098
FT DISULFID 3100 3109
FT DISULFID 3116 3127
FT DISULFID 3121 3136
FT DISULFID 3138 3147
FT DISULFID 3154 3165
FT DISULFID 3182 3274
FT DISULFID 3250 3266
FT DISULFID 3281 3324
FT DISULFID 3310 3337
FT CARBOHYD 57 57
FT CARBOHYD 331 331
FT CARBOHYD 352 352
FT CARBOHYD 817 817
FT CARBOHYD 965 965
FT CARBOHYD 1017 1017
FT CARBOHYD 1333 1333
FT CARBOHYD 1393 1393
FT CARBOHYD 1437 1437
FT CARBOHYD 1463 1463
FT CARBOHYD 1653 1653
FT CARBOHYD 1974 1974
FT CARBOHYD 2045 2045
FT CARBOHYD 2074 2074
FT CARBOHYD 2103 2103
FT CARBOHYD 2263 2263
FT CARBOHYD 2290 2290
FT CARBOHYD 2356 2356
FT CARBOHYD 2623 2623
FT CARBOHYD 2641 2641
FT CARBOHYD 2919 2919
FT CARBOHYD 3052 3052
FT CARBOHYD 3354 3354
FT CARBOHYD 3364 3364
FT VARSPLIC 349 349
P -> R (in isoform V1 and isoform V3).

FT VARSPLIC 350 1336 /FTid=VSP 003078.
FT VARSPLIC 1337 3074 /FTid=VSP 003079.
FT VARSPLIC 350 3074 /FTid=VSP 003080.
FT VARSPLIC 350 3074 /FTid=VSP 003081.
FT CONFLICT 25 25 MISSING (IN REF. 2).
FT CONFLICT 51 51 MISSING (IN REF. 2).
FT CONFLICT 89 89 N -> D (IN REF. 2).
FT CONFLICT 96 96 Q -> D (IN REF. 2).
FT CONFLICT 346 346 C -> R (IN REF. 2).
SQ SEQUENCE 3381 AA; 369984 MW; F09716FA7778D459 CRC64;
Query Match 15.8%; Score 174.5; DB 1; Length 3381;
Best Local Similarity 25.5%; Pred. No. 4.4e-07;
Matches 59; Conservative 26; Mismatches 71; Indels 75; Gaps 11;
Qy 15 QRPCYKVIYFHDTSRRLNFEFEAKERRDGGQLVLSIESEDEQKLIKFIENLLPSDGDGDF- 73
Db 3162 QGQCYK--YF---AHRRTWDAARECRLOGAHLTSILSHEEQMFVNRV-----GHDYQ 3209
Qy 74 WIGLRRREKQSNSTACQDLVAVTDGSIQFRNWTYVDEP-----SCGSEVCVMYHQPSAP 129
Db 3210 WIGL-----NDKMFEDHFRWTDGSTLQYENWRPNQPDSPFSTGEDCWIWHENG-- 3259
Qy 130 AGIGGPFYMFQWDDRCNMKNPFICKYS---DEKPAVPSRAEGE----- 170
Db 3260 -----QWNDVPCNYHLTYTCKGTGTVACQPPVVENAKTFGKMKPRYEINSLIRYHC 3310
Qy 171 -----ETELTT-----PVL-----PEETQEEDAKKTFKESREAAALN 201
Db 3311 KDGFIQRHLPTIRCLNGRWAMPKITCLNPSAYQRTYSKKYFKNSSSAKDN 3361
RESULT 7
PGCN_RAT ID_PGCN_RAT STANDARD; PRT; 1257 AA.
AC P55067;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Neurocan core protein precursor (Chondroitin sulfate proteoglycan 3)
DE (245 kDa early postnatal core glycoprotein) [Contains: 150 kDa adult
DE core glycoprotein].
GN CSPG3 OR NCAN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=Sprague-Dawley; TISSUE=Brain;
RX MEDLINE=92406907; PubMed=1326557;
RA Rauch U., Karthikeyan L., Maurel P., Margolis R.U., Margolis R.K.;
RT "Cloning and primary structure of neurocan, a developmentally
RT regulated, aggregating chondroitin sulfate proteoglycan of brain.";
RL J. Biol. Chem. 267:19536-19547(1992).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=94230574; PubMed=7513709;
RA Friedlander D.R., Milev P., Karthikeyan L., Margolis R.K.,
RA Margolis R.U., Grumet M.;
RT "The neuronal chondroitin sulfate proteoglycan neurocan binds to the
RT neural cell adhesion molecules Ng-CAM/L1/NILE and N-CAM, and inhibits
RT neuronal adhesion and neurite outgrowth.";
RL J. Cell Biol. 125:669-680(1994).
CC -!- FUNCTION: May modulate neuronal adhesion and neurite growth during
CC development by binding to neural cell adhesion molecules (NG-CAM
CC and N-CAM). Chondroitin sulfate proteoglycan; binds to hyaluronic
CC acid.
CC -!- TISSUE SPECIFICITY: EARLY POSTNATAL AND ADULT BRAIN; NOT EXPRESSED
CC IN KIDNEY, LUNG, LIVER AND MUSCLE.

[4]
SEQUENCE OF 2535-2738 FROM N.A.
STRAIN=Sprague-Dawley; TISSUE=Lung;
Blomberg L.A., Chan W.-Y., Clerch L., Massaro D.;
"Molecular cloning and characterization of two developmentally
regulated genes in rat lung."
Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
-!- FUNCTION: May play a role in intercellular signaling and in
connecting cells with the extracellular matrix. May take part in
the regulation of cell motility, growth and differentiation. Binds
hyaluronic acid.
-!- SUBUNIT: Interacts with FBLN1 (By similarity).
-!- SUBCELLULAR LOCATION: Secreted; extracellular matrix.
-!- ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=3;
Comment=Additional isoforms seem to exist;
Name=V0;
IsoId=Q9ERB4-1; Sequence=Displayed;
Name=V3;
IsoId=Q9ERB4-2; Sequence=VSP_003091;
Name=Vint;
IsoId=Q9ERB4-3; Sequence=VSP_003092;
TISSUE SPECIFICITY: In kidney is expressed in the papillary area,
but not in glomeruli.
-!- DEVELOPMENTAL STAGE: Disappears after the cartilage development
(By similarity).
-!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
-!- SIMILARITY: Contains 2 link domains.
-!- SIMILARITY: Contains 2 EGF-like domains.
-!- SIMILARITY: Contains 1 C-type lectin family domain.
-!- SIMILARITY: Contains 1 Sushi (SCR) domain.
-!- SIMILARITY: BELONGS TO THE AGGRECAN/VERSICAN PROTEOGLYCAN FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; AF062402; AAC40166.1; -.
EMBL; U75306; AAB51125.1; -.
EMBL; AF084544; AAD48544.1; -.
EMBL; AF072892; AAC26116.1; -.
EMBL; AY007691; AAG16631.1; -.
HSSP; P01132; 1EPG.
InterPro; IPR000152; Asx_hydroxyl.
InterPro; IPR000742; EGF_2.
InterPro; IPR001881; EGF_Ca.
InterPro; IPR006209; EGF_like.
InterPro; IPR007110; Ig-Like.
InterPro; IPR003599; Ig.
InterPro; IPR001304; Lectin_C.
InterPro; IPR000538; Link.
InterPro; IPR000436; Sushi_SCR_CCP.
PRINTS; PR01265; LINKMODULE.
PRINTS; PR00356; ANTIFREEZEII.
ProDom; PD000918; Link; 2.
SMART; SM00032; CCP; 1.
SMART; SM00034; CLECT; 1.
SMART; SM00179; EGF_CA; 1.
SMART; SM00409; IG_1.
SMART; SM00445; LINK; 2.
PROSITE; PS00010; ASX_HYDROXYL; 1.
PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
PROSITE; PS00022; EGF_1; 2.
PROSITE; PS01186; EGF_2; 1.
PROSITE; PS01187; EGF_CA; 1.
PROSITE; PS50835; IG_LINK; 1.
PROSITE; PS01241; LINK; 2.
Glycoprotein; Proteoglycan; Lectin; Extracellular matrix; Sushi;

KW Signal; Repeat; EGF-like domain; Calcium; Immunoglobulin domain;
KW Hyaluronic acid; Alternative splicing.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 2738 VERSICAN CORE PROTEIN.
FT NON CONS 348 349
FT DOMAIN 21 146 IG-LIKE V-TYPE.
FT DOMAIN 167 244 LINK 1.
FT DOMAIN 265 346 LINK 2.
FT DOMAIN <349 GAG-ALPHA
(GLUCOSAMINOGLYCAN ATTACHMENT DOMAIN).
FT DOMAIN 696 GAG-BETA.
FT DOMAIN 2431 EGF-LIKE 1.
FT DOMAIN 2467 EGF-LIKE 2, CALCIUM-BINDING.
FT DOMAIN 2469 2505 C-TYPE LECTIN.
FT DOMAIN 2518 2632 SUSHI.
FT DOMAIN 2637 2695 BY SIMILARITY.
FT DISULFID 44 130 BY SIMILARITY.
FT DISULFID 172 243 BY SIMILARITY.
FT DISULFID 196 217 BY SIMILARITY.
FT DISULFID 270 345 BY SIMILARITY.
FT DISULFID 294 315 BY SIMILARITY.
FT DISULFID 2435 2446 BY SIMILARITY.
FT DISULFID 2440 2455 BY SIMILARITY.
FT DISULFID 2457 2466 BY SIMILARITY.
FT DISULFID 2473 2484 BY SIMILARITY.
FT DISULFID 2478 2493 BY SIMILARITY.
FT DISULFID 2495 2504 BY SIMILARITY.
FT DISULFID 2511 2522 BY SIMILARITY.
FT DISULFID 2539 2631 BY SIMILARITY.
FT DISULFID 2607 2623 BY SIMILARITY.
FT DISULFID 2638 2681 BY SIMILARITY.
FT DISULFID 2667 2694 BY SIMILARITY.
FT CARBOHYD 57 57 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 330 330 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 692 692 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 758 758 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 802 802 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 805 805 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1257 1257 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1435 1435 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1633 1633 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1660 1660 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1684 1684 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1738 1738 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1848 1848 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2004 2004 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2409 2409 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2711 2711 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2721 2721 N-LINKED (GLCNAC. .) (POTENTIAL).
FT VARSPLIC 349 2431 Missing (in isoform V3).
FT VARSPLIC 2697 2738 /FTID=VSP_003091.
FT VARSPLIC 2738 PSAYQRTYSKKYLKNSSVVDNSINTSKHEHRSRRWQETR
FT R -> RKWSFRKNGQPCFNKY (in isoform Vint).
FT /FTID=VSP_003092.
FT AEREC -> NSARG (IN REP. 4).
FT CONFLICT 2535 2539
FT SEQUENCE 2738 AA; 300004 MW; 12CA626D58BD8C6A CRC64;
Query Match 15.7%; Score 174; DB 1; Length 2738;
Best Local Similarity 28.5%; Pred. No. 3.8e-07;
Matches 47; Conservative 23; Mismatches 55; Indels 40; Gaps 8;
QY 15 QRPCYKVIYFHDTSRRLNPFEEAKEACRRDGGQLVSISEDEQKLIKFIENLLPSDGF- 73
Db 2519 QGQCYK--YF---AHRRTWDAARECRLOGAHLTSLISHEEQMFVNRV-----GHDIQ 2566
QY 74 WIGLRRREKQSNSTACQDLYAWTDGSIQPRNYYVDEP-----SCGSEVCVMYHQPSAP 129
Db 2567 WIGL-----NDKMFEDHDFRWTGDSALQYENWRNQPDSPFSAGEDCVIWHENG-- 2616
QY 130 AGIGGPMFQWNRDRCNMKNPFICKYS-----DEKPAVPSREAEGE 170
Db 2617 -----QWNVPCNYHLTYTCKKGTGTVACGQPPVVENAKTFGK 2652

RA Paulus W., Baur I., Dours-Zimmermann M.T., Zimmermann D.R.;
RT "Differential expression of versican isoforms in brain tumors.";
RL J. Neuropathol. Exp. Neurol. 55:528-533(1996).
CC -|- FUNCTION: May play a role in intercellular signaling and in
CC connecting cells with the extracellular matrix. May take part in
CC the regulation of cell motility, growth and differentiation. Binds
CC hyaluronic acid.
CC
CC -|- SUBUNIT: Interacts with FBLN1 (By similarity).
CC -|- SUBCELLULAR LOCATION: Secreted; extracellular matrix.
CC
CC -|- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=5;
CC Comment=Additional isoforms seem to exist;
CC
CC Name=V0;
CC IsoId=P13611-1; Sequence=Displayed;
CC
CC Name=V1;
CC IsoId=P13611-2; Sequence=VSP_003082, VSP_003083;
CC
CC Name=V2;
CC IsoId=P13611-3; Sequence=VSP_003084;
CC
CC Name=V3;
CC IsoId=P13611-4; Sequence=VSP_003082, VSP_003085;
CC
CC Name=Vint;
CC IsoId=P13611-5; Sequence=VSP_003086;
CC
CC -|- TISSUE SPECIFICITY: Cerebral white matter. V0 and V1 is expressed
CC in normal brain, gliomas, medulloblastomas, schwannomas,
CC neurofibromas, and meningiomas; v2 is restricted to normal brain
CC and gliomas; v3 is found in all these tissues except
CC medulloblastomas.
CC
CC -|- DEVELOPMENTAL STAGE: Disappears after the cartilage development.
CC
CC -|- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC
CC -|- SIMILARITY: Contains 2 link domains.
CC
CC -|- SIMILARITY: Contains 1 EGF-like domains.
CC
CC -|- SIMILARITY: Contains 1 C-type lectin family domain.
CC
CC -|- SIMILARITY: Contains 1 Sushi (SCR) domain.
CC
CC -|- SIMILARITY: BELONGS TO THE AGGRECAN/VERSICAN PROTEOGLYCAN FAMILY.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U16306; AAA65018.1; -;
DR EMBL; X15998; CAA34128.1; -;
DR EMBL; S52488; AAB24878.1; -;
DR EMBL; U26555; AAA67565.1; -;
DR EMBL; D32039; BAA06801.1; -;
DR EMBL; J02814; AAA36437.1; -;
DR EMBL; AF084545; AAD48545.1; -;
DR PIR; S06014; A60979.
DR HSSP; P01132; 1EGF.
DR Genew; HGNC:2464; CSPG2.
DR MIM; 118661; -;
DR GO; GO:0005578; C:extracellular matrix; TAS.
DR GO; GO:0005204; F:chondroitin sulfate proteoglycan; TAS.
DR GO; GO:0005540; F:hyaluronic acid binding activity; TAS.
DR GO; GO:0008037; P:cell recognition; TAS.
DR GO; GO:0007275; P:development; TAS.
DR InterPro; IPR000152; Asx hydroxyl.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR000538; Link.
DR InterPro; IPR000436; Sushi_SCR_CCP.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00059; lectin_c; 1.
DR Pfam; PF00084; sushi; 1.
DR Pfam; PF00193; Xlink; 2.

DR ProDom; PD000918; Link; 2.
DR SMART; SM00032; CCP; 1.
DR SMART; SM00034; CLECT; 1.
DR SMART; SM00179; EGF_CA; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00445; LINK; 2.
DR PROSITE; PS00010; ASX HYDROXYL; 1.
DR PROSITE; PS00615; C_TYPE_LLECTIN_1; 1.
DR PROSITE; PS50041; C_TYPE_LLECTIN_2; 1.
DR PROSITE; PS00022; EGF_1_2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01187; EGF_CA; 1.
DR PROSITE; PS50835; IG LIKE; 1.
DR PROSITE; PS01241; LINK; 2.
KW Glycoprotein; Proteoglycan; Lectin; Extracellular matrix; Sushi;
KW Signal; Repeat; EGF-like domain; Calcium; Immunoglobulin domain;
KW Hyaluronic acid; Alternative splicing.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 3396 VERSICAN CORE PROTEIN.
FT DOMAIN 21 146 IG-LIKE V-TYPE.
FT DOMAIN 167 244 LINK 1.
FT DOMAIN 265 346 LINK 2.
FT DOMAIN 348 1335 GAG-ALPHA
(GLUCOSAMINOGLYCAN ATTACHMENT DOMAIN).
FT DOMAIN 1336 3089 GAG-BETA.
FT DOMAIN 3089 3125 EGF-LIKE 1.
FT DOMAIN 3127 3163 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 3176 3290 C-TYPE LECTIN.
FT DOMAIN 3295 3353 SUSHI.
FT DISULFID 44 130 BY SIMILARITY.
FT DISULFID 172 243 BY SIMILARITY.
FT DISULFID 196 217 BY SIMILARITY.
FT DISULFID 270 345 BY SIMILARITY.
FT DISULFID 294 315 BY SIMILARITY.
FT DISULFID 3093 3104 BY SIMILARITY.
FT DISULFID 3098 3113 BY SIMILARITY.
FT DISULFID 3115 3124 BY SIMILARITY.
FT DISULFID 3131 3142 BY SIMILARITY.
FT DISULFID 3136 3151 BY SIMILARITY.
FT DISULFID 3153 3162 BY SIMILARITY.
FT DISULFID 3169 3180 BY SIMILARITY.
FT DISULFID 3197 3289 BY SIMILARITY.
FT DISULFID 3265 3281 BY SIMILARITY.
FT DISULFID 3296 3339 BY SIMILARITY.
FT DISULFID 3325 3352 BY SIMILARITY.
FT CARBOHYD 57 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 330 N-LINKED (GLCNAC. .).
FT CARBOHYD 615 N-LINKED (GLCNAC. .) (POTENTIAL).

Query Match 15.7%; Score 174; DB 1; Length 3396;
Best Local Similarity 28.5%; Pred. No. 4.9e-07;
Matches 47; Conservative 23; Mismatches 55; Indels 40; Gaps 8;

QY 15 QRPCYKVIYFHDTSRRRLNFEAKEACRRDGGQLVSISEDEQKLIKFIENLLPSDGF- 73
Db 3177 QGQCYK--YF---AHRRTWDAARECRLOGAHLTSILSHEEQMFVNRV-----GHDYQ 3224

QY 74 WIGLRRREKQSNSTACQDLYAWTDGSIQFRNYYVDEP---SCGSEVCVVMYHQPSAP 129
Db 3225 WIGL-----NDKMFEDHFRWTDGSTLQYENWRPNQPDSPFSFSSAGEDCVVIWHENG-- 3274

QY 130 AGIGPYMFQWDDRCNMKNFICKYS-----DEKPAVPSREAEGE 170
Db 3275 -----QWNDVPCNYHLTYTCKKGTACGQPPVVENAKTFGK 3310

RESULT 11
PGCV_CHICK
ID_PGCV_CHICK STANDARD; PRT; 3562 AA.
AC Q90953; Q90945;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Versican core protein precursor (Large fibroblast proteoglycan)
DE (Chondroitin sulfate proteoglycan core protein 2) (PG-M).
GN CPBG2.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS V0 AND V1).
RC STRAIN=White leghorn; TISSUE=Limb bud;
RX MEDLINE=93300846; PubMed=8314802;
RA Shinomura T., Nishida Y., Ito K., Kimata K.;
RT "cDNA cloning of PG-M, a large chondroitin sulfate proteoglycan
RT expressed during chondrogenesis in chick limb buds. Alternative
RT spliced multiforms of PG-M and their relationships to versican.";
RL J. Biol. Chem. 268:14461-14469(1993).
CC -!- FUNCTION: May play a role in intercellular signaling and in
CC connecting cells with the extracellular matrix. May take part in
CC the regulation of cell motility, growth and differentiation. Binds
CC hyaluronic acid.
CC -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Comment=Additional isoforms seem to exist;
CC Name=V0;
CC IsoId=Q90953-1; Sequence=Displayed;
CC Name=V1;
CC IsoId=Q90953-2; Sequence=VSP_003093;
CC TISSUE SPECIFICITY: Prechondrogenic condensation area of
CC developing limb buds.
CC -!- DEVELOPMENTAL STAGE: Disappears after the cartilage development
CC (By similarity).
CC -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -!- SIMILARITY: Contains 2 link domains.
CC -!- SIMILARITY: Contains 2 EGF-like domains.
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC -!- SIMILARITY: Contains 1 Sushi (SCR) domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X60226; CAA42787.1; -;
DR EMBL; D13542; BAA02742.1; -;
DR PIR; A47171; A47171.
DR HSSP; P00740; LEDM.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR000538; Link.
DR InterPro; IPR000436; Sushi_SCR_CCP.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF00059; lectin_c; 1.
DR Pfam; PF00084; sushi; 1.
DR Pfam; PF00193; Xlink; 2.
DR PRINTS; PR01265; LINKMODULE.
DR ProDom; PD000918; Link; 2.
DR SMART; SM00032; CCP; 1.
DR SMART; SM00034; CLECT; 1.
DR SMART; SM00179; EGF_CA; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00445; LINK; 2.

DR PROSITE; PS00010; ASX_HYDROXYL; 1.
DR PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR PROSITE; PS00041; C_TYPE_LECTIN_2; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01187; EGF_CA; 1.
DR PROSITE; PS050835; IG_LIKE; 1.
DR PROSITE; PS01241; LINK; 2.
KW Glycoprotein; Proteoglycan; Lectin; Extracellular matrix; Sushi;
KW Signal; Repeat; EGF-like domain; Calcium; Immunoglobulin domain;
KW Hyaluronic acid; Alternative splicing.
FT SIGNAL 1 26 POTENTIAL.
FT CHAIN 27 3562 VERSICAN CORE PROTEIN.
FT DOMAIN 27 143 IG-LIKE V-TYPE.
FT DOMAIN 166 243 LINK 1.
FT DOMAIN 264 345 LINK 2.
FT DOMAIN 3254 3290 EGF-LIKE 1.
FT DOMAIN 3292 3328 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 3341 3455 C-TYPE LECTIN.
FT DOMAIN 3460 3518 SUSHI.
FT DISULFID 44 129 BY SIMILARITY.
FT DISULFID 171 242 BY SIMILARITY.
FT DISULFID 195 216 BY SIMILARITY.
FT DISULFID 269 344 BY SIMILARITY.
FT DISULFID 293 314 BY SIMILARITY.
FT DISULFID 3258 3269 BY SIMILARITY.
FT DISULFID 3263 3278 BY SIMILARITY.
FT DISULFID 3280 3289 BY SIMILARITY.
FT DISULFID 3296 3307 BY SIMILARITY.
FT DISULFID 3301 3316 BY SIMILARITY.
FT DISULFID 3318 3327 BY SIMILARITY.
FT DISULFID 3334 3345 BY SIMILARITY.
FT DISULFID 3362 3454 BY SIMILARITY.
FT DISULFID 3430 3446 BY SIMILARITY.
FT DISULFID 3461 3504 BY SIMILARITY.
FT DISULFID 3490 3517 BY SIMILARITY.
FT CARBOHYD 163 163 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 235 235 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 329 329 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 529 529 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 709 709 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 948 948 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1409 1409 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1479 1479 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1523 1523 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1530 1530 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1625 1625 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1751 1751 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1988 1988 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2088 2088 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2089 2089 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2507 2507 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2642 2642 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2679 2679 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2748 2748 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2762 2762 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 3069 3069 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 3194 3194 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 3232 3232 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 3545 3545 N-LINKED (GLCNAC. .) (POTENTIAL).
FT VARSPLIC 485 1411 Missing (in isoform V1).
FT SEQUENCE 3562 AA; 388078 MW; 9BC566E88C1602D2 CRC64;
SQ

Query Match 15.5%; Score 171; DB 1; Length 3562;
Best Local Similarity 28.5%; Pred. No. 9.7e-07;
Matches 47; Conservative 23; Mismatches 55; Indels 40; Gaps 8;

QY 15 QRPCKVIYFHDTSRRRLNFEAEACRRDGGQLVSISEDEQKLEKFIENLLPSDGF- 73
| | | | | : : : : : | | | | | : : : : :
Db 3342 QGQCYK--YF---AHRRTWDTABRECRLOGAHLTSILSHEEQVFNRI-----GHDYQ 3389
| | | | | : : : : : | | | | | : : : : :
QY 74 WIGLRRREKQSNSTACQDLYAWTDGSIQFRNWTYVDEP-----SCGSEVCVMYHQSAP 129

Db 3390 WIGL-----NDKMFERDFWTDGSPLOQYENWRPNQDPDSFFSAGEDCVIIWHENG-- 3439

QY 130 AGIGPYMFQWDDRCNMKNFICKYS-----DEKPAVPSREAEGE 170

Db 3440 -----QWNVPCNYHLTYTCKKGTVACGQPPVVENAKTFGK 3475

RESULT 12

CD93_RAT

ID_CD93_RAT STANDARD; PRT; 643 AA.

AC Q9ET61; Q9JIZ6;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Complement component C1q receptor precursor (Complement component 1, q

DE subcomponent, receptor 1) (C1qRp) (C1qR(p)) (C1q/MBL/SPA receptor)

DE (CD93 antigen) (Cell surface antigen AA4).

GN C1QR1 OR CD93 OR C1QRP.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=PVG; TISSUE=Natural killer cells;

RX MEDLINE=20507883; PubMed=11093152;

RA Dean Y.D., McGreal E.P., Akatsu H., Gasque P.;

RT "Molecular and cellular properties of the rat AA4 antigen, a C-type

RT lectin-like receptor with structural homology to thrombomodulin.";

RL J. Biol. Chem. 275:34382-34392(2000).

CC -!- FUNCTION: Receptor (or element of a larger receptor complex) for

CC C1q, mannose-binding lectin (MBL2) and pulmonary surfactant

CC protein A (SPA). May mediate the enhancement of phagocytosis in

CC monocytes and macrophages upon interaction with soluble defense

CC collagens. May play a role in intercellular adhesion.

CC -!- SUBCELLULAR LOCATION: Type I membrane protein.

CC -!- TISSUE SPECIFICITY: Widely expressed. Highly expressed in lung and

CC heart. Expressed at lower level in brain, thymus, liver, spleen,

CC intestine, kidney, adrenal gland, muscle and testis. Expressed on

CC endothelial cells, platelets, undifferentiated monocytes and

CC circulating natural killer cells.

CC -!- PTM: N- and O-glycosylated (By similarity).

CC -!- SIMILARITY: Contains 1 C-type lectin family domain.

CC -!- SIMILARITY: Contains 5 EGF-like domains.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; AF136537; AAC01572.1; -.

DR EMBL; AF160978; AAF80402.1; -.

DR HSSP; P35555; 1EMN.

DR GO; GO:0016021; C:integral to membrane; ISS.

DR GO; GO:0004872; F:receptor activity; ISS.

DR GO; GO:0016337; P:cell-cell adhesion; ISS.

DR GO; GO:0042116; P:macrophage activation; ISS.

DR GO; GO:0006909; P:phagocytosis; ISS.

DR InterPro; IPR000152; Asx_hydroxyl.

DR InterPro; IPR001881; EGF_Ca.

DR InterPro; IPR006209; EGF_like.

DR InterPro; IPR001304; Lectin_C.

DR Pfam; PF00008; EGF; 5.

DR Pfam; PF00059; lectin_c; 1.

DR SMART; SM00034; CLECT; 1.

DR SMART; SM00179; EGF_CA; 3.

DR PROSITE; PS00010; ASX_HYDROXYL; 3.

DR PROSITE; PS00615; C_TYPE_LECTIN_1; FALSE_NEG.

DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.

DR PROSITE; PS01186; EGF_2; 3.

DR PROSITE; PS01187; EGF_CA; 3.

KW Cell adhesion; Receptor; Repeat; Signal; Transmembrane;

KW EGF-like domain; Lectin; Glycoprotein.

FT SIGNAL 1 23 POTENTIAL.

FT CHAIN 24 643 COMPLEMENT COMPONENT C1Q RECEPTOR.

FT DOMAIN 24 571 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 572 592 POTENTIAL.

FT DOMAIN 593 643 CYTOPLASMIC (POTENTIAL).

FT DOMAIN 31 173 C-TYPE LECTIN.

FT DOMAIN 257 298 EGF-LIKE 1.

FT DOMAIN 299 341 EGF-LIKE 2.

FT DOMAIN 342 381 EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).

FT DOMAIN 382 423 EGF-LIKE 4, CALCIUM-BINDING (POTENTIAL).

FT DOMAIN 424 462 EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).

FT DISULFID 261 272 BY SIMILARITY.

FT DISULFID 268 282 BY SIMILARITY.

FT DISULFID 284 297 BY SIMILARITY.

FT DISULFID 303 314 BY SIMILARITY.

FT DISULFID 308 325 BY SIMILARITY.

FT DISULFID 327 340 BY SIMILARITY.

FT DISULFID 346 355 BY SIMILARITY.

FT DISULFID 351 364 BY SIMILARITY.

FT DISULFID 366 380 BY SIMILARITY.

FT DISULFID 386 397 BY SIMILARITY.

FT DISULFID 393 406 BY SIMILARITY.

FT DISULFID 408 422 BY SIMILARITY.

FT DISULFID 428 437 BY SIMILARITY.

FT DISULFID 433 446 BY SIMILARITY.

FT DISULFID 448 461 BY SIMILARITY.

FT CARBOHYD 322 322 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 498 498 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CONFLICT 417 417 E -> K (IN REF. 2).

SQ SEQUENCE 643 AA; 68781 MW; 9AE4C933AD943DB6 CRC64;

Query Match 14.9%; Score 165; DB 1; Length 643;

Best Local Similarity 25.2%; Pred. No. 4e-07;

Matches 52; Conservative 35; Mismatches 73; Indels 46; Gaps 10;

QY 1 GRLLSGQP-----VCRGGTQRPCYKVIYFHDTSRLNFEEAKEACRRDGGQLVSI ESD 54

Db 15 GQLWAGAAADSEAVVCEG---TACYTAHW-----GKLSAAEAQHRCNENGNLATV KSEE 66

QY 55 EQKLI EKFIENLL-----PSD---GDFWIGLRRREKQSNSTACQDLYAWT-DGSI SQF RN 106

Db 67 EARHVQEALAQLLKTKAPSETKIGKFWIGLQREKCKTYHDLPMKGF SWVG GEDT TYSN 126

QY 107 WY-VDEPSGSEVCVVMY-----HQPSAPAGIGGPMFQWMDR C-----NMKNF 151

Db 127 WYKASKSSCISKRCVSLILDLSLKPHPSHLP-----KWHESPCGTPDAPGNSIEGP 177

QY 152 ICKYSDEKPAVPSREAEGETELTTP 177

Db 178 LCKFNFKGMCSPALGPGQLTYTTP 203

RESULT 13

PGCA_BOVIN

ID_PGCA_BOVIN STANDARD; PRT; 2364 AA.

AC P13608; P79117; Q28159;

DT 01-JAN-1990 (Rel. 13, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE AggreCAN core protein precursor (Cartilage-specific proteoglycan core

DE protein) (CSPCP).

FT DISULFID 297 318 BY SIMILARITY.
FT DISULFID 509 580 BY SIMILARITY.
FT DISULFID 533 554 BY SIMILARITY.
FT DISULFID 607 682 BY SIMILARITY.
FT DISULFID 631 652 BY SIMILARITY.
FT DISULFID 2117 2128 BY SIMILARITY.
FT DISULFID 2182 2274 BY SIMILARITY.
FT DISULFID 2250 2266 BY SIMILARITY.
FT DISULFID 2281 2324 BY SIMILARITY.
FT DISULFID 2310 2337 BY SIMILARITY.
FT CARBOHYD 126 126 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 239 239 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 333 333 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 387 387 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 611 611 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 667 667 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 2114 2150 Missing (in isoform 2).
FT /FTid=VSP 003072.
SQ SEQUENCE 2364 AA; 246359 MW; 6FF83763420C3D4C CRC64;
Query Match 14.3%; Score 158.5; DB 1; Length 2364;
Best Local Similarity 26.8%; Pred. No. 7.6e-06;
Matches 48; Conservative 29; Mismatches 59; Indels 43; Gaps 10;
Qy 7 QPVCRGG---TQPCYKVIYFHTDTRRLNFEEAKEACRRDGGQLVSI ESEDEQKLI EKF I 63
Db 2151 QKLCEGWTKFGHCYR--HFPD---RATWVDAESQCRKQSHLSIVTPEEQ----EFV 2201
Qy 64 ENLLPSDGD F-WIGLRRREEKQSNSTACQDLYAWTDGSI SFRN WYVDEP----SCGSEV 118
Db 2202 NN---NAQDYQWIGL-----NDKTIEGDFRWSHGSLQFENWRPNQPNFFATGEDC 2250
Qy 119 CVVMYHQPSAPAGIGGYPYFQWDDRCNMKNFICKYS-----DEKPAVPSREAEGETE 173
Db 2251 VVMIWHEKG-----EWNDVPCNYQLPFTCKGTACGEPVVEHARIFGQKGD 2298
RESULT 14
PGCA_HUMAN ID PGCA_HUMAN STANDARD; PRT; 2415 AA.
AC P16112; Q13650;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Aggrecan core protein precursor (Cartilage-specific proteoglycan core
DE protein) (CSPCP) (Chondroitin sulfate proteoglycan core protein 1).
GN AGC1 OR CSPG1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 3).
RC TISSUE=Chondrocytes;
RX MEDLINE=91093289; PubMed=1985970;
RA Doege K.J., Sasaki M., Kimura T., Yamada Y.;
RT "Complete coding sequence and deduced primary structure of the human
RT cartilage large aggregating proteoglycan, aggrecan. Human-specific
RT repeats, and additional alternatively spliced forms.";
RL J. Biol. Chem. 266:894-902(1991).
RN [2]
RP SEQUENCE OF 1778-2415 FROM N.A. (ISOFORM 2).
RC TISSUE=Chondrocytes;
RA Dudhia J., Hardingham T.E.;
RL Submitted (JAN-1990) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 1936-2415 FROM N.A. (ISOFORM 1).
RX MEDLINE=89380154; PubMed=2789216;
RA Baldwin C.T., Reginato A.M., Prockop D.J.;
RT "A new epidermal growth factor-like domain in the human core protein
RT for the large cartilage-specific proteoglycan. Evidence for
RT alternative splicing of the domain.";
RL J. Biol. Chem. 264:15747-15750(1989).

RN [4]
RP SEQUENCE OF 764-864 FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=95128522; PubMed=7827755;
RA Barry F.P., Neame P.J., Sasse J., Pearson D.;
RT "Length variation in the keratan sulfate domain of mammalian
RT aggrecan.";
RL Matrix Biol. 14:323-328(1994).
CC -!- FUNCTION: THIS PROTEOGLYCAN IS A MAJOR COMPONENT OF EXTRACELLULAR
CC MATRIX OF CARTILAGINOUS TISSUES. A MAJOR FUNCTION OF THIS PROTEIN
CC IS TO RESIST COMPRESSION IN CARTILAGE. IT BINDS AVIDLY TO
CC HYALURONIC ACID VIA AN AMINO-TERMINAL GLOBULAR REGION.
CC -!- SUBUNIT: Interacts with FBLN1 (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix (By
CC similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Comment=Additional isoforms seem to exist;
CC Name=1;
CC IsoId=P16112-1; Sequence=Displayed;
CC Name=2;
CC IsoId=P16112-2; Sequence=VSP_003074;
CC Name=3;
CC IsoId=P16112-3; Sequence=VSP_003074, VSP_003075;
CC -!- DOMAIN: TWO GLOBULAR DOMAINS, G1 AND G2, COMPRISE THE AMINO
CC TERMINUS OF THE PROTEOGLYCAN, WHILE ANOTHER GLOBULAR REGION, G3,
CC MAKES UP THE COOH TERMINUS. G1 CONTAINS LINK DOMAINS AND THUS
CC CONSISTS OF THREE DISULFIDE-BONDED LOOP STRUCTURES DESIGNATED AS
CC THE A, B, B' MOTIFS. G2 IS SIMILAR TO G1. THE KERATAN SULFATE (KS)
CC AND THE CHONDROITIN SULFATE (CS) ATTACHMENT DOMAINS LIE BETWEEN G2
CC AND G3.
CC -!- PTM: CONTAINS MOSTLY CHONDROITIN SULFATE, BUT ALSO KERATAN SULFATE
CC CHAINS, N-LINKED AND O-LINKED OLIGOSACCHARIDES.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -!- SIMILARITY: Contains 4 link domains.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC -!- SIMILARITY: Contains 1 Sushi (SCR) domain.
CC -!- SIMILARITY: BELONGS TO THE AGGREGAN/VERSICAN PROTEOGLYCAN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; M55172; AAA62824.1; -
DR EMBL; J05062; AAA35726.1; -
DR EMBL; X17406; CAA35463.1; -
DR EMBL; S74659; AAC60643.2; -
DR PIR; A39086; A39086.
DR HSSP; P98066; 1TSG.
DR Genew; HGNC:319; AGC1.
DR MIM; 155760; -
DR GO; GO:0005204; P:chondroitin sulfate proteoglycan; TAS.
DR InterPro; IPR002353; AntifreezeII.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR000538; Link.
DR InterPro; IPR003324; SGXXSG.
DR InterPro; IPR000436; Sushi_SCR_CCP.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF00059; lectin_c; 1.
DR Pfam; PF02339; SGXXSG; 71.
DR Pfam; PF00084; sushi; 1.
DR Pfam; PF00193; Xlink; 4.


```

CC AND THE CHONDROITIN SULFATE (CS) ATTACHMENT DOMAINS LIE BETWEEN G2
CC AND G3.
CC
CC -!- PTM: CONTAINS MOSTLY CHONDROITIN SULFATE, BUT ALSO KERATAN SULFATE
CC CHAINS, N-LINKED AND O-LINKED OLIGOSACCHARIDES (BY SIMILARITY).
CC
CC -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
CC
CC -!- SIMILARITY: CONTAINS 4 LINK DOMAINS.
CC
CC -!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC
CC -!- SIMILARITY: CONTAINS 1 SUSHI (SCR) DOMAIN.
CC
CC -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC
CC -!- SIMILARITY: BELONGS TO THE AGGREGAN/VERSICAN PROTEOGLYCAN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC -----
CC
CC EMBL; U65989; AAB06238.2; -.
CC
CC EMBL; S74662; AAC60527.1; -.
CC
CC EMBL; L07054; -; NOT_ANNOTATED_CDS.
CC
CC PIR; I46998; I46998.
CC
CC HSSP; P08709; 1BF9.
CC
CC InterPro; IPR000152; Asx_hydroxyl.
CC
CC InterPro; IPR000742; EGF_2.
CC
CC InterPro; IPR001881; EGF_Ca.
CC
CC InterPro; IPR006209; EGF_like.
CC
CC InterPro; IPR007110; Ig-like.
CC
CC InterPro; IPR003006; Ig_MHC.
CC
CC InterPro; IPR001304; Lectin_C.
CC
CC InterPro; IPR000538; Link.
CC
CC InterPro; IPR003324; SGXSG.
CC
CC InterPro; IPR000436; Sushi_SCR_CCP.
CC
CC Pfam; PF00008; EGF; 1.
CC
CC Pfam; PF00047; ig; 1.
CC
CC Pfam; PF00059; lectin_c; 1.
CC
CC Pfam; PF02339; SGXSG; 66.
CC
CC Pfam; PF00084; sushi; 1.
CC
CC Pfam; PF00193; Xlink; 4.
CC
CC PRINTS; PR01265; LINKMODULE.
CC
CC PRINTS; PR00356; ANTIFREEZEII.
CC
CC ProDom; PD000918; Link; 4.
CC
CC SMART; SM00032; CCP; 1.
CC
CC SMART; SM00034; CLECT; 1.
CC
CC SMART; SM00179; EGF_CA; 1.
CC
CC SMART; SM00445; LINK; 4.
CC
CC PROSITE; PS00010; ASX_HYDROXYL; 1.
CC
CC PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
CC
CC PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
CC
CC PROSITE; PS00022; EGF_1; 1.
CC
CC PROSITE; PS01187; EGF_CA; 1.
CC
CC PROSITE; PS50835; IG_LIKE; 1.
CC
CC PROSITE; PS00290; IG_MHC; 1.
CC
CC PROSITE; PS01241; LINK; 4.
CC
CC KW Glycoprotein; Cartilage; Proteoglycan; Lectin; Signal; Sushi;
CC EGF-like domain; Repeat; Immunoglobulin domain.
CC
CC FT SIGNAL 1 16
CC FT CHAIN 17 2333
CC FT DOMAIN 34 147
CC FT DOMAIN 170 247
CC FT DOMAIN 268 349
CC FT DOMAIN 513 590
CC FT DOMAIN 611 692
CC FT DOMAIN 2081 2117
CC FT DOMAIN 2130 2245
CC FT DOMAIN 2249 2307
CC FT DOMAIN 48 140
CC FT DOMAIN 152 247
CC FT DOMAIN 253 349
CC FT DOMAIN 495 589
CC FT DOMAIN 596 691
CC FT DOMAIN 694 816
CC KS.

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 22, 2003, 16:06:34 ; Search time 9.91785 Seconds
(without alignments)
1978.090 Million cell updates/sec

Title: US-09-887-855-2_COPY_24_227
Perfect score: 1106
Sequence: 1 GRLLSGQPVCVRGGTQPCYK.....EEDAKKTFKESREAAALNLAY 204

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues
Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	182	16.5	1456	1 A36563	mannose receptor p
2	178.5	16.1	1455	1 A48925	mannose receptor p
3	177	16.0	1268	2 S52781	neurocan - mouse
4	174.5	15.8	1643	2 T14274	versican precursor
5	174.5	15.8	3381	2 T42389	versican precursor
6	174	15.7	1257	2 S28764	neurocan precursor
7	174	15.7	2397	1 A55535	versican precursor
8	174	15.7	2409	1 A60979	versican precursor
9	171	15.5	3562	2 A47171	chondroitin sulfat
10	170.5	15.4	1479	2 T42710	mannose receptor,
11	158.5	14.3	1340	2 A39808	proteoglycan core
12	158.5	14.3	2327	2 T42630	aggrecan - bovine
13	158.5	14.3	2415	1 A39086	aggrecan precursor
14	154.5	14.0	612	2 B42755	E-selectin precurs
15	153.5	13.9	2124	2 A28452	proteoglycan core
16	152	13.7	912	2 A54423	brevican precursor
17	149.5	13.5	459	2 T24425	hypothetical prote
18	149	13.5	321	1 LNHUER	IgE Fc receptor II
19	148.5	13.4	330	2 T46256	brevican - human (
20	148	13.4	253	2 E89130	protein F52E1.2 [i
21	147.5	13.3	2132	1 A55182	aggrecan precursor
22	147	13.3	883	2 S57653	brevican precursor
23	146.5	13.2	162	1 LNR1	lectin BRA3-1 prec
24	146	13.2	2109	1 I50421	aggrecan precursor
25	145.5	13.2	742	2 JC7595	scavenger receptor
26	145	13.1	883	2 S49126	brevican precursor
27	144.5	13.1	173	2 S10548	lectin - barnacle
28	144.5	13.1	372	2 S23936	L-selectin precurs
29	144.5	13.1	404	2 A46274	HIV gp120-binding

30	143.5	13.0	129	2 JC4329	coagulation factor
31	143.5	13.0	372	1 A32375	L-selectin precurs
32	143.5	13.0	463	2 T26655	hypothetical prote
33	142.5	12.9	131	2 JC5058	bitiscetin alpha c
34	142	12.8	1487	2 S48719	phospholipase-A(2)
35	141.5	12.8	331	1 LNMSE	IgE Fc receptor, 1
36	140.5	12.7	162	1 LNR3	lectin BRA3-2 prec
37	140	12.7	370	2 S22124	L-selectin precurs
38	139	12.6	248	1 LNHUPS	pulmonary surfacta
39	139	12.6	248	1 LNHUP6	pulmonary surfacta
40	139	12.6	248	1 LNHUP1	pulmonary surfacta
41	139	12.6	283	1 LNFHLS	lectin precursor -
42	138.5	12.5	152	2 JC4690	coagulation factor
43	138.5	12.5	202	2 JC4031	tetranectin precur
44	138	12.5	280	2 T29200	hypothetical prote
45	137.5	12.4	309	1 S34198	IgE Fc receptor II

ALIGNMENTS

RESULT 1

A36563

mannose receptor precursor - human

C:Species: Homo sapiens (man)

C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C:Accession: A36563; A60926; A44255; B44255; C44255; D44255; E44255; F44255; G44255; H

R:Taylor, M.E.; Conary, J.T.; Lennartz, M.R.; Stahl, P.D.; Drickamer, K.

J. Biol. Chem. 265, 12156-12162, 1990

A:Title: Primary structure of the mannose receptor contains multiple motifs resembling

A:Reference number: A36563; MUID:90324192; PMID:2373685

A:Accession: A36563

A:Molecule type: mRNA

A:Residues: 1-1456 <TAY>

A:Cross-references: GB:J05550; NID:g188675; PIDN:AAA59868.1; PID:g188676

A:Note: parts of this sequence, including the amino end of the mature protein, were co

R:Ezekowitz, R.A.B.; Sastry, K.; Bailly, P.; Warner, A.

J. Exp. Med. 172, 1785-1794, 1990

A:Title: Molecular characterization of the human macrophage mannose receptor: demonstr

A:Reference number: A60926; MUID:91079783; PMID:2258707

A:Accession: A60926

A>Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-1333, 'T', 1335-1456 <EZE>

A:Cross-references: GB:X55635

A:Note: translation of the nucleotide sequence is incomplete

A:Note: in the authors' translation additional residues Pro-Glu-Ile are shown after 49

R:Kim, S.J.; Ruiz, N.; Bezouska, K.; Drickamer, K.

Genomics 14, 721-727, 1992

A:Title: Organization of the gene encoding the human macrophage mannose receptor (MRC1

A:Reference number: A44255; MUID:93052405; PMID:1294118

A:Accession: A44255

A>Status: preliminary; nucleic acid sequence not shown; not compared with conceptual t

A:Molecule type: DNA

A:Residues: 155-233, 'KSAL', 238-283; 346-428; 492-569; 631-714, 716-719; 783-820, 'N', 822-865

A:Note: sequence extracted from NCBI backbone (NCBIP:118415, NCBIP:118421, NCBIP:118421,

C:Genetics:

A:Gene: GDB:MRC1

A:Cross-references: GDB:133759; OMIM:153618

A:Map position: 10p13-10p13

C:Superfamily: phospholipase A2 receptor; C-type lectin homology; fibronectin type II

C:Keywords: duplication; lectin; tandem repeat; transmembrane protein

F;1-18/Domain: signal sequence #status predicted <SIG>

F;168-209/Domain: fibronectin type II repeat homology <2F1>

F;223-340/Domain: C-type lectin homology <LCH1>

F;362-486/Domain: C-type lectin homology <LCH2>

F;945-1079/Domain: C-type lectin homology <LCH3>

Query Match 16.5%; Score 182; DB 1; Length 1456;

Best Local Similarity 25.5%; Pred. No. 8.8e-08;

Matches 52; Conservative 37; Mismatches 61; Indels 54; Gaps 8;

QY

19 YKVIYFHDTSRLNFEEAKEACRRDGGQLVSIIESEDEQKLIIEKFLENLLPSDGFWIGLR 78

Db 807 YKDYQYFSKEKETMDNARAFCKRNFGLVSIQSESEKKFLWKYV-NRNDQAQSAFYIGLL 865

QY 79 RREEKQSNSTACQDLYAWTDGSIQFRNMYVDEPSCGS--EVCVVMYHQPSAPAGIGGPY 136

Db 866 ISLDKK-----FAWMDGSKVDYVSWATGEPNFANEDENCVTMY-----SNSGF---- 908

QY 137 MFQWNDRCNMKNKFICKYSDEK----PAVPSREAEGEETELTTPVLPEETQE----- 185

Db 909 ---WNDINCGYPNAPICQRHNSINATTVM-----TPMSVPSCGCKEGWNFYSN 954

QY 186 -----EDAKKTFKESREAAAL 200

Db 955 KCFKIFGFMEERKNWQEARACI 978

RESULT 2

A48925

mannose receptor precursor, macrophage - mouse

C;Species: Mus musculus (house mouse)

C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C;Accession: A48925; S21320; PC2245

R;Harris, N.; Super, M.; Rits, M.; Chang, G.; Ezekowitz, R.A.

Blood 80, 2363-2373, 1992

A;Title: Characterization of the murine macrophage mannose receptor: demonstration that

A;Reference number: A48925; MUID:93043353; PMID:1421407

A;Accession: A48925

A;Status: not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 1-1455 <HAR>

A;Experimental source: peritoneal macrophage

A;Note: sequence extracted from NCBI backbone (NCBIP:118733)

R;Harris, N.; Super, M.; Rits, M.; Chang, G.; Ezekowitz, R.B.

submitted to the EMBL Data Library, April 1992

A;Description: Characterization of the murine macrophage mannose receptor: Demonstration

on.

A;Reference number: S21320

A;Accession: S21320

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-302,'W',303-1117,'E',1119-1455 <HA2>

A;Cross-references: EMBL:211974; NID:G52997; PIDN:CAA78028.1; PID:G52998

R;Harris, N.; Peters, L.L.; Eicher, E.M.; Rits, M.; Raspberry, D.; Eichbaum, Q.G.; Super

Biochem. Biophys. Res. Commun. 198, 682-692, 1994

A;Title: The exon-intron structure and chromosomal localization of the mouse macrophage

A;Reference number: PC2245; MUID:94128116; PMID:8297379

A;Accession: PC2245

A;Molecule type: mRNA

A;Residues: 35-105 <HA3>

C;Genetics:

A;Gene: Mrcl

A;Map position: 2

C;Superfamily: phospholipase A2 receptor; C-type lectin homology; fibronectin type II re

C;Keywords: membrane protein; receptor

F;168-209/Domain: fibronectin type II repeat homology <2F9>

F;361-485/Domain: C-type lectin homology <LCH1>

F;943-1077/Domain: C-type lectin homology <LCH2>

Query Match 16.1%; Score 178.5; DB 1; Length 1455;

Best Local Similarity 25.9%; Pred. No. 1.8e-07;

Matches 53; Conservative 33; Mismatches 66; Indels 53; Gaps 8;

QY 19 YKVIYFHDTSRRLNFEAEKACRRDGGQLVSISEDEQKLEKFIENLLPSDGDGFWIGLR 78

Db 806 YKDYQYFSKEKETMDNARRECKNFGDLATIKSESEKKFLWKYI-NKNGGQSPYFIGML 864

QY 79 RREEKQSNSTACQDLYAWTDGSIQFRNMYVDEPSCGS--EVCVVMYHQPSAPAGIGGPY 136

Db 865 ISMDKK-----FIWMDGSKVDVFAWATGEPNFANDDENCVTMY-----TNSGF---- 907

QY 137 MFQWNDRCNMKNKFICK---YSDEKPAVPSREAEGEETELTTPVLPEETQE----- 185

Db 908 ---WNDINCGYPNNFICQRHNSINATAMP-----TTPTPGCKEGWHLYKNK 953

QY 186 -----EDAKKTFKESREAAALNL 202

Db 954 CFKIFGFANEEKKSWQDARQACKGL 978

RESULT 3

S52781

neurocan - mouse

C;Species: Mus musculus (house mouse)

C;Date: 19-May-1995 #sequence_revision 21-Jul-1995 #text_change 04-Feb-2000

C;Accession: S52781

R;Rauch, U.; Forsberg, N.; Kulbe, G.; Arnold-Ammer, I.; Faessler, R.

submitted to the EMBL Data Library, February 1995

A;Description: Amino acid sequence of mouse neurocan and their different e

A;Reference number: S52781

A;Accession: S52781

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-1268 <RAU>

A;Cross-references: EMBL:X84727; NID:G758629; PIDN:CAA59216.1; PID:G758630

C;Superfamily: aggrecan; C-type lectin homology; complement factor H repeat homology; E

F;176-253/Domain: link protein repeat homology <LNK1>

F;274-355/Domain: link protein repeat homology <LNK2>

F;964-995/Domain: EGF homology <EGF>

F;1040-1160/Domain: C-type lectin homology <LCH>

F;1167-1223/Domain: complement factor H repeat homology <FHD>

Query Match 16.0%; Score 177; DB 2; Length 1268;

Best Local Similarity 31.5%; Pred. No. 2.1e-07;

Matches 45; Conservative 17; Mismatches 49; Indels 32; Gaps 6;

QY 15 QRPCYKVIYFHDTSRRLNFEAEKACRRDGGQLVSISEDEQKLEKFIENLLPSDGDGFW 74

Db 1048 QGHCYR--YF---AHRRAWEDAERDCRRAGHLTSVHSPEEHKFINSF-----GHENSW 1096

QY 75 IGLRRREKQSNSTACQDLYAWTDGSIQFRNMYVDEPS---CGSEVCVVMYHQPSAPAG 131

Db 1097 IGLNDRITVERD-----FQWTDNTGLQYENWREKQPDNFFAGGEDCVVMVAHESG--- 1145

QY 132 IGGPYMFQWNDRCNMKNKFICK 154

Db 1146 -----RWNDVPCYNLPYVCK 1161

RESULT 4

T14274

versican precursor, splice form V2 - bovine

C;Species: Bos primigenius taurus (cattle)

C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 05-May-2000

C;Accession: T14274

R;Schmalfeldt, M.; Dours-Zimmermann, M.T.; Winterhalter, K.H.; Zimmermann, D.R.

J. Biol. Chem. 273, 15758-15764, 1998

A;Title: Versican V2 is a major extracellular matrix component of the mature bovine bra

A;Reference number: Z17954; MUID:98288320; PMID:9624174

A;Accession: T14274

A;Status: preliminary; translated from GB/EMBL/DBBJ

A;Molecule type: mRNA

A;Residues: 1-1643 <SCH>

A;Cross-references: EMBL:AF060458; NID:G32533303; PID:G32533304; PIDN:AAC24360.1

A;Experimental source: brain

C;Keywords: glycoprotein

F;1-20/Domain: signal sequence #status predicted <SIG>

F;21-1643/Product: versican, splice form V2 #status predicted <MAT>

F;57,331,352,817,965,1017,1333,1616,1626/Binding site: carbohydrate (Asn) (covalent) #s

Query Match 15.8%; Score 174.5; DB 2; Length 1643;

Best Local Similarity 25.5%; Pred. No. 4.7e-07;

Matches 59; Conservative 26; Mismatches 71; Indels 75; Gaps 11;

QY 15 QRPCYKVIYFHDTSRRLNFEAEKACRRDGGQLVSISEDEQKLEKFIENLLPSDGDGFW 73

Db 1424 QGQCYSK--YF---AHRRTWDAERCRQLQGAHLTILSHEEQMFVNRV-----GHDYQ 1471

RESULT 8
A60979
verbican precursor - human
N;Alternate names: chondroitin sulfate proteoglycan 2; chondroitin sulfate proteoglycan
N;Contains: glial hyaluronate-binding protein
C;Species: Homo sapiens (man)
C;Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text_change 19-Jan-2001
C;Accession: S06014; S43921; A60979; A30358; A29348; A45131; I54179
R;Zimmermann, D.R.; Ruoslahti, E.
EMBO J. 8, 2975-2981, 1989
A;Title: Multiple domains of the large fibroblast proteoglycan, versican.
A;Reference number: S06014; MUID:90059882; PMID:2583089
A;Accession: S06014
A;Molecule type: mRNA
A;Residues: 1-2409 <ZIM>
A;Cross-references: GB:X15998; NID:g37662; PIDN:CAA34128.1; PID:g37663
R;Yao, L.Y.; Moody, C.; Schoenherr, E.; Wight, T.N.; Sandell, L.J.
Matrix Biol. 14, 213-225, 1994
A;Title: Identification of the proteoglycan versican in aorta and smooth muscle cells by
A;Reference number: S43921; MUID:95005762; PMID:7921538
A;Accession: S43921
A;Molecule type: mRNA
A;Residues: 208-440;1094-1385;1910-2246 <YAO>
R;Bignami, A.; Lane, W.S.; Andrews, D.; Dahl, D.
Brain Res. Bull. 22, 67-70, 1989
A;Title: Structural similarity of hyaluronate binding proteins in brain and cartilage.
A;Reference number: A60979; MUID:89229983; PMID:2469524
A;Accession: A60979
A;Molecule type: protein
A;Residues: 171-210;289-303 <BIG>
R;Perides, G.; Lane, W.S.; Andrews, D.; Dahl, D.; Bignami, A.
J. Biol. Chem. 264, 5981-5987, 1989
A;Title: Isolation and partial characterization of a glial hyaluronate-binding protein.
A;Reference number: A30358; MUID:89174663; PMID:2466833
A;Accession: A30358
A;Molecule type: protein
A;Residues: 24-50;80-87,'D',89-119;128-155;167-218;229-259,'IR','261-268;277-283,'G',285-
R;Krusius, T.; Gehlsen, K.R.; Ruoslahti, E.
J. Biol. Chem. 262, 13120-13125, 1987
A;Title: A fibroblast chondroitin sulfate proteoglycan core protein contains lectin-like
A;Reference number: A29348; MUID:88007514; PMID:2820964
A;Accession: A29348
A;Molecule type: mRNA
A;Residues: 1725,'V',1727-2409 <KRU>
A;Cross-references: GB:J02814
R;Perides, G.; Rahemtulla, F.; Lane, W.S.; Asher, R.A.; Bignami, A.
J. Biol. Chem. 267, 23883-23887, 1992
A;Title: Isolation of a large aggregating proteoglycan from human brain.
A;Reference number: A45131; MUID:93054750; PMID:1429726
A;Contents: brain
A;Accession: A45131
A;Molecule type: protein
A;Residues: 21-22,'X',24-37 <PE2>
A;Experimental source: brain
A;Note: sequence extracted from NCBI backbone (NCBIP:118884)
R;Iozzo, R.V.; Naso, M.F.; Cannizzaro, L.A.; Wasmuth, J.J.; McPherson, J.D.
Genomics 14, 845-851, 1992
A;Title: Mapping of the versican proteoglycan gene (CSPG2) to the long arm of human chro
A;Reference number: I54179; MUID:93122792; PMID:1478664
A;Accession: I54179
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 251-347 <RES>
A;Cross-references: GB:S52488; NID:G263313; PIDN:AAB24878.1; PID:g263314
C;Genetics:
A;Gene: GDB:CSPG2
A;Cross-references: GDB:127873; OMIM:118661
A;Map position: 5q12-5q14
C;Superfamily: versican; C-type lectin homology; complement factor H repeat homology; Eg
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-2409/Product: proteoglycan 24K core protein #status predicted <MAT>
F;167-244/Domain: link protein repeat homology <LNK1>

F;265-346/Domain: link protein repeat homology <LNK2>
F;559-1654/Domain: chondroitin sulfate attachment #status predicted <GAG>
F;2106-2137/Domain: EGF homology <EG1>
F;2144-2175/Domain: EGF homology <EG2>
F;2182-2302/Domain: C-type lectin homology <LCH>
F;2309-2365/Domain: complement factor H repeat homology <FHD>

Query Match 15.7%; Score 174; DB 1; Length 2409;
Best Local Similarity 28.5%; Pred. No. 8e-07;
Matches 47; Conservative 23; Mismatches 55; Indels 40; Gaps 8;

Qy 15 QRPCYKVIYFHDTSRRRLNFEAKEACRRDGGQLVSI ESEDEQKLI EKFIENLLPSD GDF- 73
| | | | | : : : : : | | | | | : : : : :
Db 2190 QGQCYK--YF---AHRRTWDAARECRLQGAHLT SILSHEEQVFNRV-----GHDYQ 2237

Qy 74 WIGLRRREEKQSNSTACQDLYAWTDGSI SQFRNMYVDEP----SCGSEVCVVMYHQPSAP 129
| | | | | : : : : : | | | | | : : : : :
Db 2238 WIGL-----NDKMFEDFRWTDGSTLQYENWRPNQPD SFFSAGEDCVVIWHENG-- 2287

Qy 130 AGIGGPMFQWDDRCNMKNFNICKYS----DEKPAVPSREAEGE 170
| | | | | : : : : : | | | | | : : : : :
Db 2288 -----QWNDVPCNYHLTYTCKKGTGTVACGPPVVENAKTFGK 2323

RESULT 9
A47171
chondroitin sulfate proteoglycan PG-M core protein - chicken
C;Species: Gallus gallus (chicken)
C;Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 21-Jul-2000
C;Accession: A47171
R;Shinomura, T.; Nishida, Y.; Ito, K.; Kimata, K.
J. Biol. Chem. 268, 14461-14469, 1993
A;Title: cDNA cloning of PG-M, a large chondroitin sulfate proteoglycan expressed durin

A;Reference number: A47171; MUID:93300846; PMID:8314802
A;Accession: A47171
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-3562 <SHI>
A;Cross-references: GB:D13542; NID:g391643; PIDN:BAA02742.1; PID:g391644
A;Experimental source: stage 22-23 developing limb buds
A;Note: sequence extracted from NCBI backbone (NCBIN:134456, NCBIP:134457)
C;Superfamily: chicken chondroitin sulfate proteoglycan PG-M core protein; C-type lecti
F;166-243/Domain: link protein repeat homology <LNK1>
F;264-345/Domain: link protein repeat homology <LNK2>
F;3258-3289/Domain: EGF homology <EGF1>
F;3296-3327/Domain: EGF homology <EGF>
F;3334-3454/Domain: C-type lectin homology <LCH>
F;3461-3517/Domain: complement factor H repeat homology <FHD>

Query Match 15.5%; Score 171; DB 2; Length 3562;
Best Local Similarity 28.5%; Pred. No. 2.3e-06;
Matches 47; Conservative 23; Mismatches 55; Indels 40; Gaps 8;

Qy 15 QRPCYKVIYFHDTSRRRLNFEAKEACRRDGGQLVSI ESEDEQKLI EKFIENLLPSD GDF- 73
| | | | | : : : : : | | | | | : : : : :
Db 3342 QGQCYK--YF---AHRRTWDAARECRLQGAHLT SILSHEEQVFNRI-----GHDYQ 3389

Qy 74 WIGLRRREEKQSNSTACQDLYAWTDGSI SQFRNMYVDEP----SCGSEVCVVMYHQPSAP 129
| | | | | : : : : : | | | | | : : : : :
Db 3390 WIGL-----NDKMFEDFRWTDGSTLQYENWRPNQPD SFFSAGEDCVVIWHENG-- 3439

Qy 130 AGIGGPMFQWDDRCNMKNFNICKYS----DEKPAVPSREAEGE 170
| | | | | : : : : : | | | | | : : : : :
Db 3440 -----QWNDVPCNYHLTYTCKKGTGTVACGPPVVENAKTFGK 3475

RESULT 10
T42710
mannose receptor, macrophage - mouse
N;Alternate names: lambda lectin; phospholipase A2 receptor
C;Species: Mus musculus (house mouse)
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jun-2000

C;Accession: T42710
R;Wu, K.; Yuan, J.; Lasky, L.A.
J. Biol. Chem. 271, 21323-21330, 1996
A;Title: Characterization of a novel member of the macrophage mannose receptor type C 1e
A;Reference number: 222235; MUID:96355501; PMID:8702911
A;Accession: T42710
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1479 <WUK>
A;Cross-references: EMBL:U56734; NID:G1336073; PID:G1336074; PIDN:AAC52729.1
C;Superfamily: phospholipase A2 receptor; C-type lectin homology; fibronectin type II re
C;Keywords: membrane protein; receptor
F;186-227/Domain: fibronectin type II repeat homology <2FR>

Query Match 15.4%; Score 170.5; DB 2; Length 1479;
Best Local Similarity 31.4%; Pred. No. 9.3e-07;
Matches 48; Conservative 22; Mismatches 48; Indels 35; Gaps 7;

QY 5 SGQPVCRGGTQPCYKVIYFHDTSRRLNFEAEACRRDGGQLVSIIESEDEQKLEKFI 64
Db 384 SWQPF-----QGHCYRL-----QAEKRSWQESKRACLRGGGDLLSIHSMAELEFITKIK 433

QY 65 NLLPSDGFWMGLRRREKQSNSTACQDLYAWTDGSIQFRNWWYVDEPS---CGSEVCVV 121
Db 434 QEVE---ELWIGL-----NDLKLQMNFEWSDGSLVSFTHWHPFEPNFRDLSLEDCVT 482

QY 122 MYHQPSAPAGIGGPPYMFQWNDRCNMKNPFICK 154
Db 483 IW----GPEG-----RWNDSPCNQSLPSICK 504

RESULT 11
A39808
proteoglycan core protein, cartilage - bovine (fragments)
N;Alternate names: aggrecan; aggregating cartilage proteoglycan
C;Species: Bos primigenius taurus (cattle)
C;Date: 20-Mar-1992 #sequence revision 23-Mar-1995 #text change 13-Aug-1999
C;Accession: A34234; A27752; A39808; A27751; E29164; B27751; E27751; F27
R;Antonsson, P.; Heinegard, D.; Oldberg, A.
J. Biol. Chem. 264, 16170-16173, 1989
A;Title: The keratan sulfate-enriched region of bovine cartilage proteoglycan consists o
A;Reference number: A34234; MUID:89380219; PMID:2528543
A;Accession: A34234
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 128-621 <ANT>
A;Cross-references: GB:J05028
R;Oldberg, A.; Antonsson, P.; Heinegard, D.
Biochem. J. 243, 255-259, 1987
A;Title: The partial amino acid sequence of bovine cartilage proteoglycan, deduced from a
A;Reference number: A27752; MUID:87270630; PMID:3111460
A;Accession: A27752
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 622-1340 <OLD>
R;Sandy, J.D.; Boynton, R.E.; Flannery, C.R.
J. Biol. Chem. 266, 8198-8205, 1991
A;Title: Analysis of the catabolism of aggrecan in cartilage explants by quantitation of
A;Reference number: A39808; MUID:91217051; PMID:2022637
A;Accession: A39808
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-28;59-82;131-137,'QSET',142-149;196-207;226-249;1137-1143;1252-1267;1274-1
R;Perin, J.P.; Bonnet, F.; Jolles, P.
FEBS Lett. 206, 73-77, 1986
A;Title: Structural relationship between link proteins and proteoglycan monomers.
A;Reference number: A27751; MUID:87005253; PMID:3530809
A;Accession: A27751
A;Molecule type: protein
A;Residues: 29-58;74-130;174-175,'A',177-204;208-225 <PER>
R;Perin, J.P.; Bonnet, F.; Jolles, J.; Jolles, P.
FEBS Lett. 176, 37-42, 1984
A;Title: Sequence data concerning the protein core of the cartilage proteoglycan monomer

A;Reference number: A91327; MUID:85027710; PMID:6489519
A;Accession: E29164
A;Molecule type: protein
A;Residues: 1230-1249 <PE2>
C;Superfamily: aggrecan; C-type lectin homology; complement factor H repeat homology;
C;Keywords: glycoprotein
F;8-28/Domain: link protein repeat homology (fragment) <LNK1>
F;29-58/Domain: link protein repeat homology (fragment) <LNK2>
F;80-146/Domain: link protein repeat homology (fragments) <LNK3>
F;167-248/Domain: link protein repeat homology <LNK4>
F;1130-1250/Domain: C-type lectin homology <LCH>
F;1257-1313/Domain: complement factor H repeat homology <FHD>

Query Match 14.3%; Score 158.5; DB 2; Length 1340;
Best Local Similarity 26.8%; Pred. No. 9.6e-06;
Matches 48; Conservative 29; Mismatches 59; Indels 43; Gaps 10;

QY 7 QPVCRRGG---TORPCYKVIYFHDTSRRLNFEAEACRRDGGQLVSIIESEDEQKLEKFI 63
Db 1127 QKLCIEGWTKFQGHCHYR--HFPD---RATWVDAESQCRKQKQSHLSSIVTPEEQ----EFV 1177

QY 64 ENLLPSDGDGF-WIGLRRREKQSNSTACQDLYAWTDGSIQFRNWWYVDEP----SCGSEV 118
Db 1178 NN---NAQDYQWIGL-----NDKTIEGDFRWSDGHSLOQFENWRPNQDNFFATGEDC 1226

QY 119 CVVMYHQPSAPAGIGGPPYMFQWNDRCNMKNPFICKYS-----DEKPAVPSREAEGETE 173
Db 1227 VVMIWHEKG-----EWNDVPCNYQLPFTCKGTACGEPVVEHARIFGQKGD 1274

RESULT 12
T42630
aggrecan - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000
C;Accession: T42630
R;Hering, T.M.; Kollar, J.; Huynh, T.D.
submitted to the EMBL Data Library, September 1996
A;Description: Complete coding sequence of bovine aggrecan: comparative structural ana
A;Reference number: 222182
A;Accession: T42630
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-2327 <HER>
A;Cross-references: EMBL:U76615; NID:G1730259; PID:G1730260; PIDN:AAB38524.1
A;Experimental source: articular chondrocytes
C;Superfamily: aggrecan; C-type lectin homology; complement factor H repeat homology;
C;Keywords: cartilage; chondroitin sulfate proteoglycan; extracellular matrix; glycopr

Query Match 14.3%; Score 158.5; DB 2; Length 2327;
Best Local Similarity 26.8%; Pred. No. 1.8e-05;
Matches 48; Conservative 29; Mismatches 59; Indels 43; Gaps 10;

QY 7 QPVCRRGG---TORPCYKVIYFHDTSRRLNFEAEACRRDGGQLVSIIESEDEQKLEKFI 63
Db 2114 QKLCIEGWTKFQGHCHYR--HFPD---RATWVDAESQCRKQKQSHLSSIVTPEEQ----EFV 2164

QY 64 ENLLPSDGDGF-WIGLRRREKQSNSTACQDLYAWTDGSIQFRNWWYVDEP----SCGSEV 118
Db 2165 NN---NAQDYQWIGL-----NDKTIEGDFRWSDGHSLOQFENWRPNQDNFFATGEDC 2213

QY 119 CVVMYHQPSAPAGIGGPPYMFQWNDRCNMKNPFICKYS-----DEKPAVPSREAEGETE 173
Db 2214 VVMIWHEKG-----EWNDVPCNYQLPFTCKGTACGEPVVEHARIFGQKGD 2261

RESULT 13
A39086
aggrecan precursor, cartilage long splice form [validated] - human
N;Alternate names: chondroitin sulfate proteoglycan 1; large aggregating proteoglycan;
N;Contains: aggrecan cartilage short splice form
C;Species: Homo sapiens (man)
C;Date: 10-Sep-1999 #sequence_revision 01-Dec-2000 #text_change 08-Dec-2000

F;1-21/Domain: signal sequence #status predicted <SIG>
F;12-138/Domain: C-type lectin homology <LCH>
F;22-612/Product: p-selectin #status predicted <MAT>
F;143-174/Domain: EGF homology <EGF>
F;180-238/Domain: complement factor H repeat homology <FH1>
F;243-300/Domain: complement factor H repeat homology <FH2>
F;305-363/Domain: complement factor H repeat homology <FH3>
F;368-426/Domain: complement factor H repeat homology <FH4>
F;431-489/Domain: complement factor H repeat homology <FH5>
F;494-548/Domain: complement factor H repeat homology <FH6>
F;25,391,528/Binding site: carbohydrate (Asn) (covalent) #s

Query Match	14.0%	Score 154.5;	DB 2;	Length 612;
Best Local Similarity	27.9%	Pred. No. 8.7e-06;		
Matches 38; Conservative	28;	Mismatches 45;	Indels 25;	Gaps 5;

Qy	23 YFHDTSRRLLNFEEAEACRRDGGQLVSEIEDEQEKLIEKFLENLLPSDGFWIGURRREE	82
	: : : : : : : : : : :	
D_b	23 YYNASSELMTYDEASAYCQRDYTHLVAIQNKEE---INVLNSNLKHSPSYWVGIRK---	76

QY 83 QKSNSTACQDLYAWTDGS---ISQFRNWWYDEPS--CGSEVCVVMVHQPSAPAGIGPYM 137

QY	138	FQW	NDDRC	NMKN	NFIC	153
				:		:
DB	125	--W	NDERC	NKKK	LALC	138

RESULT 15

A28452
proteoglycan core protein precursor, cartilage - rat
N:Alternate names: aggrecan
C:Species: Rattus norvegicus (Norway rat)
C:Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 21-Jan-2000
R:Accession: A92623; A23835; A28453; A28095; A28452
C:Doerge, K.; Sasaki, M.; Horigan, E.; Hassell, J.R.; Yamada, Y.
J. Biol. Chem. 262, 17757-17767, 1987
A:Title: Complete primary structure of the rat cartilage proteoglycan core protein deduced from complementary DNA sequence
A:Reference number: A92623; MUID:88087070; PMID:3693370
A:Accession: A92623
A:Molecule type: mRNA
A:Residues: 1-2124 <DOE>
R:Doerge, K.; Sasaki, M.; Horigan, E.; Hassell, J.R.; Yamada, Y.
J. Biol. Chem. 263, 10040a, 1988
A:Reference number: A30069
A:Contents: annotation; revision to residue 698
R:Doerge, K.; Fernandez, P.; Hassell, J.R.; Sasaki, M.; Yamada, Y.
J. Biol. Chem. 261, 8108-8111, 1986
A:Title: Partial cDNA sequence encoding a globular domain at the C terminus of the rat core protein
A:Reference number: A23835; MUID:86250698; PMID:2424893
A:Accession: A23835
A:Molecule type: mRNA
A:Residues: 1856-2124 <DO2>
A:Cross-references: GB:M13518; NID:g206104; PIDN:AAA41836.1; PID:g206105
R:Neame, P.J.; Christner, J.E.; Baker, J.R.
J. Biol. Chem. 262, 17768-17778, 1987
A:Title: Cartilage proteoglycan aggregates. The link protein and proteoglycan amino-terminal domain
A:Reference number: A28453; MUID:88087071; PMID:3693371
A:Accession: A28453
A:Molecule type: protein
A:Residues: 20-37, 'W', 39-60, 'E', 62-64, 'X', 66-69; 70-83; 84, 89-148, 'L', 150-238, 'S', 240, 'A',
C:Superfamily: aggrecan; C-type lectin homology; complement factor H repeat homology; Ed
C:Keywords: glycoprotein
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-2124/Product: proteoglycan core protein #status predicted <MAT>
F:44-135/Domain: immunoglobulin homology <IMW>
F:170-247/Domain: link protein repeat homology <LNK1>
F:268-349/Domain: link protein repeat homology <LNK2>
F:504-581/Domain: link protein repeat homology <LNK3>
F:602-683/Domain: link protein repeat homology <LNK4>
F:1914-2034/Domain: C-type lectin homology <LCH>
F:2041-2097/Domain: complement factor H repeat homology <FHD>

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 22, 2003, 16:07:10 ; Search time 19.4937 Seconds
(without alignments)
1954.412 Million cell updates/sec

Title: US-09-887-855-2 COPY 24 227

Perfect score: 1106

Sequence: 1 GRLLSGQPVCRGGTQRPCYK.....EEDAKKTFKESREALNLAY 204

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 696363 seqs, 186758610 residues

Total number of hits satisfying chosen parameters: 696363

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:★

```

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pcp.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pcp.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pcp.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pcp.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pcp.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pcp.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pcp.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pcp.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pcp.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pcp.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pcp.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pcp.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pcp.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pcp.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pcp.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pcp.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pcp.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pcp.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %			DB	ID	Description
		Match	Length				
1	1106	100.0	206	9	US-09-887-855-5	Sequence 5, Appli	
2	1106	100.0	374	9	US-09-887-855-2	Sequence 2, Appli	
3	1106	100.0	374	12	US-10-094-749-2090	Sequence 2090, Ap	
4	1106	100.0	374	15	US-10-149-819-15	Sequence 15, Appl	
5	1102	99.6	374	12	US-10-094-749-2142	Sequence 2142, Ap	
6	1092	98.7	382	10	US-09-909-320-137	Sequence 137, App	
7	1092	98.7	382	10	US-09-909-088B-137	Sequence 137, App	
8	1092	98.7	382	10	US-09-905-291A-137	Sequence 137, App	
9	1092	98.7	382	10	US-09-902-853-137	Sequence 137, App	
10	1092	98.7	382	10	US-09-907-824-137	Sequence 137, App	
11	1092	98.7	382	10	US-09-907-841-137	Sequence 137, App	
12	1092	98.7	382	11	US-09-904-011-137	Sequence 137, App	
13	1092	98.7	382	11	US-09-906-742-137	Sequence 137, App	
14	1092	98.7	382	11	US-09-906-838-137	Sequence 137, App	
15	1092	98.7	382	11	US-09-907-613-137	Sequence 137, App	

16	1092	98.7	382	11	US-09-907-942-137	Sequence 137, App
17	1092	98.7	382	11	US-09-904-859-137	Sequence 137, App
18	1092	98.7	382	11	US-09-909-204-137	Sequence 137, App
19	1092	98.7	382	11	US-09-904-820-137	Sequence 137, App
20	1092	98.7	382	11	US-09-904-786-137	Sequence 137, App
21	1092	98.7	382	11	US-09-906-646-137	Sequence 137, App
22	1092	98.7	382	11	US-09-906-700-137	Sequence 137, App
23	1092	98.7	382	11	US-09-903-786-137	Sequence 137, App
24	1092	98.7	382	11	US-09-902-903-137	Sequence 137, App
25	1092	98.7	382	11	US-09-903-749A-137	Sequence 137, App
26	1092	98.7	382	11	US-09-904-119-137	Sequence 137, App
27	1092	98.7	382	11	US-09-904-956-137	Sequence 137, App
28	1092	98.7	382	11	US-09-902-736-137	Sequence 137, App
29	1092	98.7	382	11	US-09-907-794-137	Sequence 137, App
30	1092	98.7	382	11	US-09-903-943-137	Sequence 137, App
31	1092	98.7	382	11	US-09-904-462-137	Sequence 137, App
32	1092	98.7	382	11	US-09-907-925-137	Sequence 137, App
33	1092	98.7	382	11	US-09-902-692-137	Sequence 137, App
34	1092	98.7	382	11	US-09-903-520-137	Sequence 137, App
35	1092	98.7	382	11	US-09-905-056-137	Sequence 137, App
36	1092	98.7	382	11	US-09-909-064-137	Sequence 137, App
37	1092	98.7	382	11	US-09-904-553-137	Sequence 137, App
38	1092	98.7	382	11	US-09-905-381-137	Sequence 137, App
39	1092	98.7	382	11	US-09-905-088-137	Sequence 137, App
40	1092	98.7	382	11	US-09-907-575-137	Sequence 137, App
41	1092	98.7	382	11	US-09-905-075-137	Sequence 137, App
42	1092	98.7	382	11	US-09-902-759-137	Sequence 137, App
43	1092	98.7	382	11	US-09-902-634-137	Sequence 137, App
44	1092	98.7	382	11	US-09-902-713-137	Sequence 137, App
45	1092	98.7	382	11	US-09-907-979-137	Sequence 137, App

ALIGNMENTS

```

RESULT 1
US-09-887-855-5
; Sequence 5, Application US/09887855
; Patent No. US20020058310A1
; GENERAL INFORMATION:
; APPLICANT: Immunex Corporation
; APPLICANT: Anderson, Dirk M
; TITLE OF INVENTION: LECTIN SS3939 DNA AND POLYPEPTIDES
; FILE REFERENCE: 2883-US
; CURRENT APPLICATION NUMBER: US/09/887,855
; CURRENT FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 206
; TYPE: PRT
; ORGANISM: Homo sapiens
;
US-09-887-855-5

```

Query Match	100.0%;	Score 1106;	DB 9;	Length 206;
Best Local Similarity	100.0%;	Pred. NO. 6.7e-104;		
Matches 204;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	GRLLSGQPVCRGGTQPCYKVIFYHDT	SRRLNFEFAKEACRRDGGQLV	SI ESEDEQKLI E 60
Db	3	GRLLSGQPVCRGGTQPCYKVIFYHDT	SRRLNFEFAKEACRRDGGQLV	SI ESEDEQKLI E 62
QY	61	KFIENLLPSDGFWIGLRRREEKQSN	STACQDLYAWTDGSI	SQFRNWWYDEPSCGSEVCV 120
Db	63	KFIENLLPSDGFWIGLRRREEKQSN	STACQDLYAWTDGSI	SQFRNWWYDEPSCGSEVCV 122
QY	121	VMYHQPSAPAGIGGPFYMFQWDDRC	NMKNFICKYSDEKPAVPSREAE	GETELTPVLP 180
Db	123	VMYHQPSAPAGIGGPFYMFQWDDRC	NMKNFICKYSDEKPAVPSREAE	GETELTPVLP 182
QY	181	EETQEEADAKKTFKESREAA	NLAY 204	
Db	183	EETQEEADAKKTFKESREAA	NLAY 206	

```

RESULT 2
US-09-887-855-2
; Sequence 2, Application US/09887855
; Patent No. US20020058310A1
; GENERAL INFORMATION:
; APPLICANT: Immunex Corporation
; APPLICANT: Anderson, Dirk M
; TITLE OF INVENTION: LECTIN SS3939 DNA AND POLYPEPTIDES
; FILE REFERENCE: 2883-US
; CURRENT APPLICATION NUMBER: US/09/887,855
; CURRENT FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-887-855-2

Query Match      100.0%; Score 1106; DB 9; Length 374;
Best Local Similarity 100.0%; Pred. No. 1.5e-103;
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GRLLSGQPVCRGGTQRPCVKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI ESEDEQKLI E 60
Db      24 GRLLSGQPVCRGGTQRPCVKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI ESEDEQKLI E 83

QY      61 KFIENLLPSDGFWIGLRRRREKQSNSTACQDLYAWTDGSI SQFRNWWYVDEPSCGSEVCV 120
Db      84 KFIENLLPSDGFWIGLRRRREKQSNSTACQDLYAWTDGSI SQFRNWWYVDEPSCGSEVCV 143

QY      121 VMYHQPSAPAGIGGPPYMFQWDDRCNMKNFICKYSDEKPAVPSRAEGEETELTTPVLP 180
Db      144 VMYHQPSAPAGIGGPPYMFQWDDRCNMKNFICKYSDEKPAVPSRAEGEETELTTPVLP 203

QY      181 EETQEEDAKKTFKESREAAALNLAY 204
Db      204 EETQEEDAKKTFKESREAAALNLAY 227

```

```

; APPLICANT: SHAH, Purvi
; APPLICANT: LAL, Preeti
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: BURFORD, Neil
; TITLE OF INVENTION: EXTRACELLULAR MATRIX AND CELL ADHESION MOLECULES
; FILE REFERENCE: PF-0760 PCT
; CURRENT APPLICATION NUMBER: US/10/149,819
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/172,852; 60/172,354
; PRIOR FILING DATE: 1999-12-10; 1999-12-16
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PERL Program
; SEQ ID NO 15
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030044913A1 31434111CD1
US-10-149-819-15

```

```

; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2090
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-094-749-2090

Query Match          100.0%; Score 1106; DB 12; Length 374;
Best Local Similarity 100.0%; Pred. No. 1.5e-103;
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0

Qy      1  GRLLSGQPVCRRGGTQPCYKVIFYFHDTSRRLLNFEEAKEACRRDGGQLVSIESEDEQKLI 60
        |||||
Db      24 GRLLSGQPVCRRGGTQPCYKVIFYFHDTSRRLLNFEEAKEACRRDGGQLVSIESEDEQKLI 83

Qy      61 KFIENLLPSDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNWWYDEPSCGSEVCV 120
        |||||
Db      84 KFIENLLPSDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNWWYDEPSCGSEVCV 143

Qy      121 VMYHQPSAPAGIGGPPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGEETELTTPVLP 180
        |||||
Db      144 VMYHQPSAPAGIGGPPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGEETELTTPVLP 203

Qy      181 EETQEEDAKKTFKESREAAALNLAY 204
        |||||
Db      204 EETQEEDAKKTFKESREAAALNLAY 227

```

```

RESULT 4
US-10-149-819-15
; Sequence 15, Application US/10149819
; Publication No. US20030044913A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: YUE, Henry
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: TANG, Y. Tom
; APPLICANT: PATTERSON, Chandra
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: SHAH, Purvi
; APPLICANT: LAL, Preeti
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: BURFORD, Neil
; TITLE OF INVENTION: EXTRACELLULAR MATRIX AND CELL ADHESION MOLECULES
; FILE REFERENCE: PF-0760 PCT
; CURRENT APPLICATION NUMBER: US/10/149,819
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/172,852; 60/172,354
; PRIOR FILING DATE: 1999-12-10; 1999-12-16
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PERL Program
; SEQ ID NO 15
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030044913A1 3143411CD1
US-10-149-819-15

```

QY 121 VMYHQPAPAGIGGPPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGETELTTPVLP 180
Db 144 VMYHQPAPAGIGGPPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGETELTTPVLP 203
QY 181 EETQEDAKTKFKESREAAALNLAY 204
Db 204 EETQEDAKTKFKESREAAALNLAY 227

RESULT 5

US-10-094-749-2142
; Sequence 2142, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOUYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2142
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-094-749-2142

Query Match 99.6%; Score 1102; DB 12; Length 374;
Best Local Similarity 99.5%; Pred. No. 3.8e-103;
Matches 203; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GRLLSGQPVCRCGGTQPCVKVIYFHDTSRRLNFEAEAKACRRDGGQLVSI ESEDEQKLI E 60
Db 24 GRLLSGQPVCRCGGTQPCVKVIYFHDTSRRLNFEAEAKACRRDGGQLVSI ESEDEQKLI E 83
QY 61 KFIEILLPSDGDGFWIGLRRREEKQSNSTACQDLYAWTDGSI SQFRNWWYVDEPSCGSEVCV 120
Db 84 KFIEILLPSDGDGFWIGLRRREEKQSNSTACQDLYAWTDGSI SQFRNWWYVDEPSCGSEVCV 143
QY 121 VMYHQPAPAGIGGPPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGETELTTPVLP 180
Db 144 VMYHQPAPAGIGGPPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGETELTTPVLP 203
QY 181 EETQEDAKTKFKESREAAALNLAY 204
Db 204 EETQEDAKTKFKESREAAALNLAY 227

RESULT 6

US-09-909-320-137
; Sequence 137, Application US/09909320

Patent No. US20020132240A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kijavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/909,320
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-909-320-137

Query Match 98.7%; Score 1092; DB 10; Length 382;
Best Local Similarity 96.2%; Pred. No. 4e-102;
Matches 204; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

Qy 1 GRLLS-----GQPVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIES 52
Db 24 GRLLSASDLRLRGQPVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIES 83

Qy 53 EDEQKLIKFIENLLPSDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNMYVDEP 112
Db 84 EDEQKLIKFIENLLPSDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNMYVDEP 143

Qy 113 SCGSEVCVMYHQPAPAGIGGYPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGET 172
Db 144 SCGSEVCVMYHQPAPAGIGGYPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGET 203

Qy 173 ELTTPVLPEETOEDAKKTFKESREAAALNLAY 204
Db 204 ELTTPVLPEETOEDAKKTFKESREAAALNLAY 235

RESULT 7
US-09-909-088B-137
; Sequence 137, Application US/09909088B
; Patent No. US20020146709A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/909,088B
; CURRENT FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089

; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-909-088B-137

Query Match 98.7%; Score 1092; DB 10; Length 382;
Best Local Similarity 96.2%; Pred. No. 4e-102;
Matches 204; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

Qy 1 GRLLS-----GQPVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIES 52
Db 24 GRLLSASDLRLRGQPVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIES 83

Qy 53 EDEQKLIKFIENLLPSDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNMYVDEP 112
Db 84 EDEQKLIKFIENLLPSDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNMYVDEP 143

Qy 113 SCGSEVCVMYHQPAPAGIGGYPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGET 172
Db 144 SCGSEVCVMYHQPAPAGIGGYPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGET 203

Qy 173 ELTTPVLPEETOEDAKKTFKESREAAALNLAY 204
Db 204 ELTTPVLPEETOEDAKKTFKESREAAALNLAY 235

RESULT 8
US-09-905-291A-137
; Sequence 137, Application US/09905291A
; Patent No. US20020160374A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel


```

; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/905,291A
; CURRENT FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-905-291A-137

Query Match      98.7%; Score 1092; DB 10; Length 382;
Best Local Similarity 96.2%; Pred. No. 4e-102;
Matches 204; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY      1 GRLLS-----GQVCRGGTQPCYKVIYFHTSRRLNFEAKEACRRDGGQLVSIES 52
      |||||
Db      24 GRLLSASDLRLGGQPVCRGGTQPCYKVIYFHTSRRLNFEAKEACRRDGGQLVSIES 83

QY      53 EDEQKLIKFIENLLPSDGDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNMYVDEP 112
      |||||
Db      84 EDEQKLIKFIENLLPSDGDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNMYVDEP 143

QY      113 SCGSEVCVMYHQPSAPAGIGGPFQWDDRRCNMKNFNICKYSDEKPAVPSREAAGEET 172
      |||||
Db      144 SCGSEVCVMYHQPSAPAGIGGPFQWDDRRCNMKNFNICKYSDEKPAVPSREAAGEET 203

QY      173 ELTTPVLPEETQEDAKKTFKESREAAALNLAY 204
      |||||
Db      204 ELTTPVLPEETQEDAKKTFKESREAAALNLAY 235

RESULT 9
US-09-902-853-137
; Sequence 137, Application US/09902853
; Publication No. US20020192659A1
```

```

; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/902,853
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US/09/665,350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-902-853-137

Query Match      98.7%; Score 1092; DB 10; Length 382;
```

```
Best Local Similarity 96.2%; Pred. No. 4e-102;
Matches 204; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 GRLLS-----GQVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIES 52
Db 24 GRLLSASDLRLGGQPVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIES 83
QY 53 EDEQKLIKFIENLLPSDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNWWYVDEP 112
Db 84 EDEQKLIKFIENLLPSDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNWWYVDEP 143
QY 113 SCGSEVCVMYHQPSAPAGIGGYPMFQWNDRCNMKNNFICKYSDEKPAVPSREAEGEET 172
Db 144 SCGSEVCVMYHQPSAPAGIGGYPMFQWNDRCNMKNNFICKYSDEKPAVPSREAEGEET 203
QY 173 ELTTPVLPEETQEEADAKKTFKESREAALNLAY 204
Db 204 ELTTPVLPEETQEEADAKKTFKESREAALNLAY 235
```

RESULT 10

```
US-09-907-824-137
; Sequence 137, Application US/09907824
; Publication No. US20020197671A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/907,824
; CURRENT FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: 09/665,350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
```

```
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-907-824-137

Query Match 98.7%; Score 1092; DB 10; Length 382;
Best Local Similarity 96.2%; Pred. No. 4e-102;
Matches 204; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 GRLLS-----GQVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIES 52
Db 24 GRLLSASDLRLGGQPVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIES 83
QY 53 EDEQKLIKFIENLLPSDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNWWYVDEP 112
Db 84 EDEQKLIKFIENLLPSDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNWWYVDEP 143
QY 113 SCGSEVCVMYHQPSAPAGIGGYPMFQWNDRCNMKNNFICKYSDEKPAVPSREAEGEET 172
Db 144 SCGSEVCVMYHQPSAPAGIGGYPMFQWNDRCNMKNNFICKYSDEKPAVPSREAEGEET 203
QY 173 ELTTPVLPEETQEEADAKKTFKESREAALNLAY 204
Db 204 ELTTPVLPEETQEEADAKKTFKESREAALNLAY 235
```

RESULT 11

```
US-09-907-841-137
; Sequence 137, Application US/09907841
; Publication No. US20020198366A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
```

; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/907,841
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-907-841-137

Query Match 98.7%; Score 1092; DB 10; Length 382;
Best Local Similarity 96.2%; Pred. No. 4e-102;
Matches 204; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 GRLLS-----GQPVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIES 52
Db 24 GRLLSASDLRLGGQPVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIES 83

QY 53 EDEQKLIIEFIENLLPSDGDGFWIGLRRRREKQSNSTACQDLYAWTDGSIQFRNWWYVDEP 112
Db 84 EDEQKLIIEFIENLLPSDGDGFWIGLRRRREKQSNSTACQDLYAWTDGSIQFRNWWYVDEP 143

QY 113 SCGSEVCVMYHQPSAPAGIGGPFQWMDRCKNKNFICKYSDEKPAVPSREAEGET 172
Db 144 SCGSEVCVMYHQPSAPAGIGGPFQWMDRCKNKNFICKYSDEKPAVPSREAEGET 203

QY 173 ELTPVLPEETQEEADAKTKFKESREAAALNLAY 204
Db 204 ELTPVLPEETQEEADAKTKFKESREAAALNLAY 235

RESULT 12
US-09-904-011-137
; Sequence 137, Application US/09904011
; Publication No. US20030003530A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/904,011
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 09/665,350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-904-011-137

Query Match 98.7%; Score 1092; DB 11; Length 382;
Best Local Similarity 96.2%; Pred. No. 4e-102;
Matches 204; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 GRLLS-----GQPVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIES 52
Db 24 GRLLSASDLRLGGQPVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIES 83

QY 53 EDEQKLIIEFIENLLPSDGDGFWIGLRRRREKQSNSTACQDLYAWTDGSIQFRNWWYVDEP 112
Db 84 EDEQKLIIEFIENLLPSDGDGFWIGLRRRREKQSNSTACQDLYAWTDGSIQFRNWWYVDEP 143

QY 113 SCGSEVCVMYHQPSAPAGIGGPPYMFQWNDRCNMKNPFICKYSDEKPAVPSREAEGET 172
Db 144 SCGSEVCVMYHQPSAPAGIGGPPYMFQWNDRCNMKNPFICKYSDEKPAVPSREAEGET 203
QY 173 ELTTPVLPEETOEDAKKTFKESREAAALNLAY 204
Db 204 ELTTPVLPEETOEDAKKTFKESREAAALNLAY 235

RESULT 13

US-09-906-742-137
; Sequence 137, Application US/09906742
; Publication No. US20030023054A1
; GENERAL INFORMATION:

; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.

; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

; TITLE OF INVENTION: Acids Encoding the Same

; FILE REFERENCE: 10466-14

; CURRENT APPLICATION NUMBER: US/09/906,742

; CURRENT FILING DATE: 2001-07-16

; PRIOR APPLICATION NUMBER: 09/665,350

; PRIOR FILING DATE: 2000-09-18

; PRIOR APPLICATION NUMBER: PCT/US00/04414

; PRIOR FILING DATE: 2000-02-22

; PRIOR APPLICATION NUMBER: US 60/143,048

; PRIOR FILING DATE: 1999-07-07

; PRIOR APPLICATION NUMBER: US 60/145,698

; PRIOR FILING DATE: 1999-07-26

; PRIOR APPLICATION NUMBER: US 60/146,222

; PRIOR FILING DATE: 1999-07-28

; PRIOR APPLICATION NUMBER: PCT/US99/20594

; PRIOR FILING DATE: 1999-09-08

; PRIOR APPLICATION NUMBER: PCT/US99/20944

; PRIOR FILING DATE: 1999-09-13

; PRIOR APPLICATION NUMBER: PCT/US99/21090

; PRIOR FILING DATE: 1999-09-15

; PRIOR APPLICATION NUMBER: PCT/US99/21547

; PRIOR FILING DATE: 1999-09-15

; PRIOR APPLICATION NUMBER: PCT/US99/23089

; PRIOR FILING DATE: 1999-10-05

; PRIOR APPLICATION NUMBER: PCT/US99/28214

; PRIOR FILING DATE: 1999-11-29

; PRIOR APPLICATION NUMBER: PCT/US99/28313

; PRIOR FILING DATE: 1999-11-30

; PRIOR APPLICATION NUMBER: PCT/US99/28564

; PRIOR FILING DATE: 1999-12-02

; PRIOR APPLICATION NUMBER: PCT/US99/28565

; PRIOR FILING DATE: 1999-12-02

; PRIOR FILING DATE: 1999-12-02

; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-906-742-137

Query Match 98.7%; Score 1092; DB 11; Length 382;

Best Local Similarity 96.2%; Pred. No. 4e-102;

Matches 204; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 GRLLS-----GQVCRGGTQPCYKVIYFHDTSRRLNFEAEKACRRDGGQLVSIES 52
Db 24 GRLLSASDLRLGGQPCVCRGGTQPCYKVIYFHDTSRRLNFEAEKACRRDGGQLVSIES 83
QY 53 EDEQKLIKFIENLLPSDGDWFGLRRREKQSNSTACQDLYAWTDGSIQFRNMYVDEP 112
Db 84 EDEQKLIKFIENLLPSDGDWFGLRRREKQSNSTACQDLYAWTDGSIQFRNMYVDEP 143
QY 113 SCGSEVCVMYHQPSAPAGIGGPPYMFQWNDRCNMKNPFICKYSDEKPAVPSREAEGET 172
Db 144 SCGSEVCVMYHQPSAPAGIGGPPYMFQWNDRCNMKNPFICKYSDEKPAVPSREAEGET 203
QY 173 ELTTPVLPEETOEDAKKTFKESREAAALNLAY 204
Db 204 ELTTPVLPEETOEDAKKTFKESREAAALNLAY 235

RESULT 14

US-09-906-838-137

; Sequence 137, Application US/09906838

; Publication No. US20030027143A1

; GENERAL INFORMATION:

; APPLICANT: Genentech, Inc.

; APPLICANT: Ashkenazi, Avi

; APPLICANT: Botstein, David

; APPLICANT: Desnoyers, Luc

; APPLICANT: Eaton, Dan L.

; APPLICANT: Ferrara, Napoleone

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Fong, Sherman

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Gerber, Hanspeter

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, A.

; APPLICANT: Godowski, Paul J.

; APPLICANT: Grimaldi, Christopher J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Hillan, Kenneth, J.

; APPLICANT: Kljavin, Ivar J.

; APPLICANT: Mather, Jennie P.

; APPLICANT: Pan, James

; APPLICANT: Paoni, Nicholas F.

; APPLICANT: Roy, Margaret Ann

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tumas, Daniel

; APPLICANT: Williams, P. Mickey

; APPLICANT: Wood, William, I.

; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

; TITLE OF INVENTION: Acids Encoding the Same

; FILE REFERENCE: 10466-14

; CURRENT APPLICATION NUMBER: US/09/906,838

; CURRENT FILING DATE: 2001-07-16

; PRIOR APPLICATION NUMBER: 09/665,350

; PRIOR FILING DATE: 2000-09-18

; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-906-838-137

Query Match 98.7%; Score 1092; DB 11; Length 382;
Best Local Similarity 96.2%; Pred. No. 4e-102;
Matches 204; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

Qy 1 GRLLS-----GQVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIES 52
Db 24 GRLLSASDLRLGGQPCVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIES 83
Qy 53 EDEQKLIKFIENLLPSDGFWIGLRRREKQSNSTACQDLYAWTDGSIQFRNWWYVDEP 112
Db 84 EDEQKLIKFIENLLPSDGFWIGLRRREKQSNSTACQDLYAWTDGSIQFRNWWYVDEP 143
Qy 113 SCGSEVCVMYHQPSAPAGIGGYPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGET 172
Db 144 SCGSEVCVMYHQPSAPAGIGGYPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGET 203
Qy 173 ELTTPVLPEETQEEADAKTKFKESREAAALNLAY 204
Db 204 ELTTPVLPEETQEEADAKTKFKESREAAALNLAY 235

RESULT 15
US-09-907-613-137
; Sequence 137, Application US/09907613
; Publication No. US20030027145A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone

; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/907,613
; PRIOR FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-907-613-137

Query Match 98.7%; Score 1092; DB 11; Length 382;
Best Local Similarity 96.2%; Pred. No. 4e-102;
Matches 204; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
Qy 1 GRLLS-----GQVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIES 52
Db 24 GRLLSASDLRLGGQPCVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIES 83

Qy	53	EDEQKLEKFIENLLPSDGD	FWIGLRRREEKQSNSTACQDLYAWTDG	SISQFRN	WYVDEP	112
Db	84	EDEQKLEKFIENLLPSDGD	FWIGLRRREEKQSNSTACQDLYAWTDG	SISQFRN	WYVDEP	143
Qy	113	SCGSEVCVVMYHQPSAPAGIGGP	YMFQWNDRCNMKNFICKYSDEKPAVPSREAE	GEET		172
Db	144	SCGSEVCVVMYHQPSAPAGIGGP	YMFQWNDRCNMKNFICKYSDEKPAVPSREAE	GEET		203
Qy	173	ELTTPVLPEETQEEDAKKTFKESREAA	LNLAY			204
Db	204	ELTTPVLPEETQEEDAKKTFKESREAA	LNLAY			235

Search completed: December 22, 2003, 16:16:47
Job time : 19.4937 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 22, 2003, 16:05:09 ; Search time 6.56161 Seconds
(without alignments)
1476.391 Million cell updates/sec

Title: US-09-887-855-2_COPY_22_227
Perfect score: 1115
Sequence: 1 ATGRLSGQPVCRGGTQRPC.....EEDAKKTFKESREAAALNLAY 206

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	574	51.5	273	1	CHOD_MOUSE
2	561.5	50.4	273	1	CHOD_HUMAN
3	182	16.3	1456	1	MANR_HUMAN
4	177	15.9	1268	1	PGCN_MOUSE
5	176	15.8	1321	1	PGCN_HUMAN
6	174.5	15.7	3381	1	PGCV_BOVIN
7	174	15.6	1257	1	PGCN_RAT
8	174	15.6	2738	1	PGCV_RAT
9	174	15.6	3358	1	PGCV_MOUSE
10	174	15.6	3396	1	PGCV_HUMAN
11	171	15.3	3562	1	PGCV_CHICK
12	165	14.8	643	1	CD93_RAT
13	158.5	14.2	2364	1	PGCA_BOVIN
14	158.5	14.2	2415	1	PGCA_HUMAN
15	155.5	13.9	2333	1	PGCA_CANFA
16	154.5	13.9	612	1	LEM2_MOUSE
17	153.5	13.8	644	1	CD93_MOUSE
18	153.5	13.8	2124	1	PGCA_RAT
19	152	13.6	912	1	PGCB_BOVIN
20	151.5	13.6	652	1	CD93_HUMAN
21	151	13.5	197	1	CLF1_HUMAN
22	149	13.4	321	1	FCE2_HUMAN
23	148.5	13.3	2132	1	PGCA_MOUSE
24	147	13.2	883	1	PGCB_MOUSE
25	146	13.1	2109	1	PGCA_CHICK
26	145	13.0	158	1	LECG_TRIST
27	145	13.0	883	1	PGCB_RAT
28	144.5	13.0	173	1	LEC2_MEGRO
29	144.5	13.0	372	1	LEM1_RAT
30	143.5	12.9	372	1	LEM1_MOUSE
31	141.5	12.7	331	1	FCE2_MOUSE
32	141.5	12.7	549	1	LEM2_RAT
33	140.5	12.6	162	1	LEC3_MEGRO
					P07439 megabalanus

34	140	12.6	370	1	LEM1_BOVIN	P98131 bos taurus
35	139	12.5	248	1	PSPA_HUMAN	P07714 homo sapien
36	139	12.5	283	1	LECA_SARPE	P05047 sarcophaga
37	138.5	12.4	152	1	IXA_TRIFL	P23806 trimeresuru
38	138.5	12.4	202	1	TETN_MOUSE	P43025 mus musculus
39	137.5	12.3	372	1	LEM1_MACMU	Q95198 macaca mula
40	137.5	12.3	372	1	LEM1_PAPHA	Q28768 papio hamad
41	136	12.2	175	1	LITH_BOVIN	P23132 bos taurus
42	135	12.1	165	1	LITI_MOUSE	P43137 mus musculus
43	134.5	12.1	132	1	ACAL_ANSAN	P83300 anser anser
44	134.5	12.1	372	1	LEM1_PONPY	Q95235 pongo pygma
45	134.5	12.1	485	1	LEM2_BOVIN	P98107 bos taurus

ALIGNMENTS

RESULT 1
CHOD_MOUSE
ID CHOD_MOUSE STANDARD; PRT; 273 AA.
AC Q9CXM0; Q8VI31;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Chondrolectin precursor (Transmembrane protein MT75).
GN CHODL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RA Weng L., Smits P., Hubner R., Wouters J., Merregaert J.;
RT "Mt75, a low expressed c-type lectin gene involving in
RT chondrogenesis";
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryonic head;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; AF311699; AAL50354.1; -.

DR PFam; PF00193; Xlink; 2.
DR PRINTS; PR01265; LINKMODULE.
DR PRINTS; PR00356; ANTIFREEZEII.
DR ProDom; PD000918; Link; 2.
DR SMART; SM00032; CCP; 1.
DR SMART; SM00034; CLECT; 1.
DR SMART; SM00179; EGF_CA; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00445; LINK; 2.
DR PROSITE; PS00010; ASX HYDROXYL; 1.
DR PROSITE; PS00615; C-TYPE LECTIN_1; 1.
DR PROSITE; PS50041; C-TYPE LECTIN_2; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01187; EGF_CA; 1.
DR PROSITE; PS50835; IG LIKE; 1.
DR PROSITE; PS01241; LINK; 2.
KW Glycoprotein; Proteoglycan; Lectin; Extracellular matrix; Sushi;
KW Signal; Repeat; EGF-like domain; Calcium; Immunoglobulin domain;
KW Hyaluronic acid; Alternative splicing.
FT SIGNAL 1 20
FT CHAIN 21 3381
FT DOMAIN 21 147
FT DOMAIN 168 245
FT DOMAIN 266 347
FT DOMAIN 349 1336
FT DOMAIN 1337 3074
FT DOMAIN 3074 3110
FT DOMAIN 3112 3148
FT DOMAIN 3161 3275
FT DOMAIN 3280 3338
FT DISULFID 44 131
FT DISULFID 173 244
FT DISULFID 197 218
FT DISULFID 271 346
FT DISULFID 295 316
FT DISULFID 3078 3089
FT DISULFID 3083 3098
FT DISULFID 3100 3109
FT DISULFID 3116 3127
FT DISULFID 3121 3136
FT DISULFID 3138 3147
FT DISULFID 3154 3165
FT DISULFID 3182 3274
FT DISULFID 3250 3266
FT DISULFID 3281 3324
FT DISULFID 3310 3337
FT CARBOHYD 57 57
FT CARBOHYD 331 331
FT CARBOHYD 352 352
FT CARBOHYD 817 817
FT CARBOHYD 965 965
FT CARBOHYD 1017 1017
FT CARBOHYD 1333 1333
FT CARBOHYD 1393 1393
FT CARBOHYD 1437 1437
FT CARBOHYD 1463 1463
FT CARBOHYD 1653 1653
FT CARBOHYD 1974 1974
FT CARBOHYD 2045 2045
FT CARBOHYD 2074 2074
FT CARBOHYD 2103 2103
FT CARBOHYD 2263 2263
FT CARBOHYD 2290 2290
FT CARBOHYD 2356 2356
FT CARBOHYD 2623 2623
FT CARBOHYD 2641 2641
FT CARBOHYD 2919 2919
FT CARBOHYD 3052 3052
FT CARBOHYD 3354 3354
FT CARBOHYD 3364 3364
FT VARSPLIC 349 349
P -> R (in isoform V1 and isoform V3).

FT VARSPLIC 350 1336 /FTid=VSP_003078.
FT Missing (in isoform V1).
FT VARSPLIC 1337 3074 /FTid=VSP_003079.
FT Missing (in isoform V2).
FT VARSPLIC 350 3074 /FTid=VSP_003080.
FT Missing (in isoform V3).
FT CONFLICT 25 25 MISSING (IN REF. 2).
FT CONFLICT 51 51 MISSING (IN REF. 2).
FT CONFLICT 89 89 N -> D (IN REF. 2).
FT CONFLICT 96 96 Q -> D (IN REF. 2).
FT CONFLICT 346 346 C -> R (IN REF. 2).
SQ SEQUENCE 3381 AA; 369984 MW; F09716FA7778D459 CRC64;
Query Match 15.7%; Score 174.5; DB 1; Length 3381;
Best Local Similarity 25.5%; Pred. No. 4.5e-07;
Matches 59; Conservative 26; Mismatches 71; Indels 75; Gaps 11;
QY 17 QRPCYKVIYFHDTSRRLLNFEEAKEACRRDGGQLVLSIESEDEQKLEKFIENLLPSDGF- 75
Db 3162 QGQCYK--YF---AHRRTWDAERECRLQGAHLTSILSHEEQMFVNRV-----GHDYQ 3209
QY 76 WIGLRREEKQSNSTACQDLYAWTDGSIQFRNWWYVDEP-----SCGSEVCVVMYHQPSAP 131
Db 3210 WIGL-----NDKMFEDHDFRTDGTSLQYENWRPNQPDSPFSTGDCVVIWHENG-- 3259
QY 132 AGIGSPYMFQWDDRCNMKNFICKYS---DEKPAVPSREAEGE----- 172
Db 3260 -----QWNDVPCNYHLTYTCKKGTACGQPPVVENAKTFGKMKPRYEINSLIRYHC 3310
QY 173 -----ETELTT-----PVL-----PEETOEDAKKTFKESREAAALN 203
Db 3311 KDGFIQRHLPTIRCLNGRWAMPKITCLNPSAYQRTYSKKYFKNSSSAKDN 3361
RESULT 7
PGCN_RAT ID_PGCN_RAT STANDARD; PRT; 1257 AA.
AC P55067;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Neurocan core protein precursor (Chondroitin sulfate proteoglycan 3)
DE (245 kDa early postnatal core glycoprotein) [Contains: 150 kDa adult
core glycoprotein].
GN CSPG3 OR NCAN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=Sprague-Dawley; TISSUE=Brain;
RX MEDLINE=92406907; PubMed=1326557;
RA Rauch U., Karthikeyan L., Maurel P., Margolis R.U., Margolis R.K.;
RT "Cloning and primary structure of neurocan, a developmentally
regulated, aggregating chondroitin sulfate proteoglycan of brain.";
RL J. Biol. Chem. 267:19536-19547(1992).
RN [2]
CHARACTERIZATION.
RX MEDLINE=94230574; PubMed=7513709;
RA Friedlander D.R., Milev P., Karthikeyan L., Margolis R.K.,
RA Margolis R.U., Grumet M.;
RT "The neuronal chondroitin sulfate proteoglycan neurocan binds to the
neural cell adhesion molecules Ng-CAM/L1/NILE and N-CAM, and inhibits
neural adhesion and neurite outgrowth.";
RL J. Cell Biol. 125:669-680(1994).
CC -!- FUNCTION: May modulate neuronal adhesion and neurite growth during
development by binding to neural cell adhesion molecules (NG-CAM
and N-CAM). Chondroitin sulfate proteoglycan; binds to hyaluronic
acid.
CC -!- TISSUE SPECIFICITY: EARLY POSTNATAL AND ADULT BRAIN; NOT EXPRESSED
IN KIDNEY, LUNG, LIVER AND MUSCLE.

CC -|- PTM: CONTAINS MOSTLY CHONDROITIN SULFATE, BUT ALSO N-LINKED AND
CC O-LINKED OLIGOSACCHARIDES (BY SIMILARITY).
CC -|- PTM: TWO ISOFORMS WERE FOUND THAT PROBABLY ARISE BY PROTEOLYTIC
CC DEGRADATION. THE LARGE ISOFORM IS PREDOMINANT IN EARLY POSTNATAL
CC BRAIN, THE SMALL ISOFORM IS FOUND IN ADULT BRAIN.
CC -|- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
CC -|- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
CC -|- SIMILARITY: CONTAINS 2 LINK DOMAINS.
CC -|- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -|- SIMILARITY: CONTAINS 1 CUSHI (SCR) DOMAIN.
CC -|- SIMILARITY: BELONGS TO THE AGGREGAN/VERSICAN PROTEOGLYCAN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; M97161; AAC37679.1; -;
CC PIR; S28764; S28764.
CC HSSP; P00740; 1EDM.
CC InterPro; IPR002353; AntifreezeII.
CC InterPro; IPR000152; Asx_hydroxyl.
CC InterPro; IPR000742; EGF_2.
CC InterPro; IPR001881; EGF_Ca.
CC InterPro; IPR006209; EGF-like.
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR003599; Ig.
CC InterPro; IPR003006; Ig_MHC.
CC InterPro; IPR001304; Lectin_C.
CC InterPro; IPR000538; Link.
CC InterPro; IPR000436; Sushi_SCR_CCP.
CC Pfam; PF00008; EGF; 2.
CC Pfam; PF00047; ig; 1.
CC Pfam; PF00059; lectin_c; 1.
CC Pfam; PF00084; sushi; 1.
CC Pfam; PF00193; xlink; 2.
CC PRINTS; PR01265; LINKMODULE.
CC PRINTS; PR00356; ANTIFREEZEII.
CC ProDom; PD000918; Link; 2.
CC SMART; SM00032; CCP; 1.
CC SMART; SM00034; CLECT; 1.
CC SMART; SM00179; EGF_CA; 1.
CC SMART; SM00409; IG; 1.
CC SMART; SM00445; LINK; 2.
CC PROSITE; PS00010; ASX_HYDROXYL; 1.
CC PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
CC PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
CC PROSITE; PS00022; EGF_1; 3.
CC PROSITE; PS01186; EGF_2; 1.
CC PROSITE; PS01187; EGF_CA; 1.
CC PROSITE; PS50835; IG_LIKE; 1.
CC PROSITE; PS01241; LINK; 2.
KW Glycoprotein; Hyaluronic acid; Proteoglycan; Immunoglobulin domain;
KW EGF-like domain; Calcium; Repeat; Lectin; Sushi; Signal.
FT SIGNAL 1 22
FT CHAIN 23 1257 NEUROCAN CORE PROTEIN.
FT CHAIN 639 1257 150 kDa ADULT CORE GLYCOPROTEIN.
FT DOMAIN 37 157 IG-LIKE V-TYPE.
FT DOMAIN 158 253 LINK 1.
FT DOMAIN 259 355 LINK 2.
FT DOMAIN 949 985 EGF-LIKE 1.
FT DOMAIN 987 1023 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 1025 1154 C-TYPE LECTIN.
FT DOMAIN 1155 1213 SUSHI.
FT DISULFID 58 139 BY SIMILARITY.
FT DISULFID 181 252 BY SIMILARITY.
FT DISULFID 205 226 BY SIMILARITY.
FT DISULFID 279 354 BY SIMILARITY.
FT DISULFID 303 324 BY SIMILARITY.
FT DISULFID 953 964 BY SIMILARITY.

FT DISULFID 958 973 BY SIMILARITY.
FT DISULFID 975 984 BY SIMILARITY.
FT DISULFID 1029 1040 BY SIMILARITY.
FT DISULFID 1057 1149 BY SIMILARITY.
FT DISULFID 1125 1141 BY SIMILARITY.
FT DISULFID 1156 1199 BY SIMILARITY.
FT DISULFID 1185 1212 BY SIMILARITY.
FT CARBOHYD 121 121 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 737 737 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 944 944 O-LINKED (XYL. . .) (CHONDROITIN SULFATE).
FT CARBOHYD 967 967 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1164 1164 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1257 AA; 135544 MW; 992B33DCFA19EB1B CRC64;

Query Match 15.6%; Score 174; DB 1; Length 1257;
Best Local Similarity 30.8%; Pred. No. 1.5e-07;
Matches 44; Conservative 18; Mismatches 49; Indels 32; Gaps 6;

QY 17 QRPCYKVIYFHTSRRLNFEAKEACRRDGGQLVSISEDEQKLIKFIENLLPSDGF 76
Db 1037 QGHCYR--YF---AHRRAWEDAERCRRRAGHLTSVHSPEHKFINSF-----GHENSW 1085

QY 77 IGLRRREKQSNSTACQDLYAWTDGSIQFRNWMYVDEPS---CGSEVCVMYHQPSAPAG 133
Db 1086 IGLNDRTVTERD-----FQWTDNTGLQYENWREKQPDNFFAGGEDCVMVVAHENG--- 1134

QY 134 IGGPYMFQWNDRCNMKNFICK 156
Db 1135 -----RWNDVPCYNLPHYVCK 1150

RESULT 8
PGCV_RAT
ID_PGCV_RAT STANDARD; PRT; 2738 AA.
AC Q9ERB4; O08592; O88564; Q9R1K4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Versican core protein precursor (Large fibroblast proteoglycan)
DE (Chondroitin sulfate proteoglycan core protein 2) (PG-M) (Glial
DE hyaluronate-binding protein) (GHAP) (Fragments).
GN CP5G2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE OF 349-2738 FROM N.A. (ISOFORM V0), SEQUENCE FROM N.A.
RP (ISOFORM V3), AND SEQUENCE OF 2657-2738 FROM N.A. (ISOFORM VINT).
RC STRAIN=Wistar Kyoto;
RX MEDLINE=99327053; PubMed=10397680;
RA Lemire J.M., Braun K.R., Maurel P., Kaplan E.D., Schwartz S.M.,
RA Wight T.N.;
RT "Versican/PG-M isoforms in vascular smooth muscle cells."
RL Arterioscler. Thromb. Vasc. Biol. 19:1630-1639(1999).
RN [2]
RP SEQUENCE OF 349-2738 FROM N.A. (ISOFORM V0).
RC STRAIN=Wistar Kyoto;
RX MEDLINE=98308094; PubMed=9642104;
RA Milev P., Maurel P., Chiba A., Mevissen M., Popp S., Yamaguchi Y.,
RA Margolis R.K., Margolis R.U.;
RT "Differential regulation of expression of hyaluronan-binding
RT proteoglycans in developing brain: aggrecan, versican, neurocan, and
RT brevican."
RL Biochem. Biophys. Res. Commun. 247:207-212(1998).
RN [3]
RP SEQUENCE OF 2421-2463 FROM N.A. (ISOFORM V0).
RC TISSUE=Kidney;
RX MEDLINE=98094159; PubMed=9434070;
RA Pyke C., Kristensen P., Ostergaard P.B., Oturai P.S., Romer J.;
RT "Proteoglycan expression in the normal rat kidney."
RL Nephron 77:461-470(1997).

RN [4] SEQUENCE OF 2535-2738 FROM N.A.
RP STRAIN=Sprague-Dawley; TISSUE=Lung;
RA Blomberg L.A.; Chan W.-Y.; Clerch L.; Massaro D.;
RT "Molecular cloning and characterization of two developmentally
RL regulated genes in rat lung."
CC Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: May play a role in intercellular signaling and in
CC connecting cells with the extracellular matrix. May take part in
CC the regulation of cell motility, growth and differentiation. Binds
CC hyaluronic acid.
CC -!- SUBUNIT: Interacts with FBLN1 (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Comment=Additional isoforms seem to exist;
CC Name=V0;
CC IsoId=Q9ERB4-1; Sequence=Displayed;
CC Name=V3;
CC IsoId=Q9ERB4-2; Sequence=VSP_003091;
CC Name=Vint;
CC IsoId=Q9ERB4-3; Sequence=VSP_003092;
CC TISSUE SPECIFICITY: In kidney is expressed in the papillary area,
CC but not in glomeruli.
CC -!- DEVELOPMENTAL STAGE: Disappears after the cartilage development
CC (By similarity).
CC -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -!- SIMILARITY: Contains 2 link domains.
CC -!- SIMILARITY: Contains 2 EGF-like domains.
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC -!- SIMILARITY: Contains 1 Sushi (SCR) domain.
CC -!- SIMILARITY: BELONGS TO THE AGGRECAN/VERSICAN PROTEOGLYCAN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF062402; AAC40166.1; -.
DR EMBL; U75306; AAB51125.1; -.
DR EMBL; AF084544; AAD48544.1; -.
DR EMBL; AF072892; AAC26116.1; -.
DR EMBL; AY007691; AAG16631.1; -.
DR HSSP; P01132; 1EPG.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR007110; Ig_Like.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR000538; Link.
DR InterPro; IPR000436; Sushi_SCR_CCP.
DR PRINTS; PR01265; LINKMODULE.
DR PRINTS; PR00356; ANTIFREEZEII.
DR ProDom; PD000918; Link; 2.
DR SMART; SM00032; CCP; 1.
DR SMART; SM00034; CLECT; 1.
DR SMART; SM00179; EGF_CA; 1.
DR SMART; SM00409; IG_1.
DR SMART; SM00445; LINK; 2.
DR PROSITE; PS00010; ASX_HYDROXYL; 1.
DR PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01187; EGF_CA; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR PROSITE; PS01241; LINK; 2.
KW Glycoprotein; Proteoglycan; Lectin; Extracellular matrix; Sushi;

KW Signal; Repeat; EGF-like domain; Calcium; Immunoglobulin domain;
KW Hyaluronic acid; Alternative splicing.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 2738 VERSICAN CORE PROTEIN.
FT NON CONS 348 349
FT DOMAIN 21 146 IG-LIKE V-TYPE.
FT DOMAIN 167 244 LINK 1.
FT DOMAIN 265 346 LINK 2.
FT DOMAIN <349 695 GAG-ALPHA
(GLUCOSAMINOGLYCAN ATTACHMENT DOMAIN).
FT DOMAIN 696 2431 GAG-BETA.
FT DOMAIN 2431 2467 EGF-LIKE 1.
FT DOMAIN 2469 2505 EGF-LIKE 2, CALCIUM-BINDING.
FT DOMAIN 2518 2632 C-TYPE LECTIN.
FT DOMAIN 2637 2695 SUSHI.
FT DISULFID 44 130 BY SIMILARITY.
FT DISULFID 172 243 BY SIMILARITY.
FT DISULFID 196 217 BY SIMILARITY.
FT DISULFID 270 345 BY SIMILARITY.
FT DISULFID 294 315 BY SIMILARITY.
FT DISULFID 2435 2446 BY SIMILARITY.
FT DISULFID 2440 2455 BY SIMILARITY.
FT DISULFID 2457 2466 BY SIMILARITY.
FT DISULFID 2473 2484 BY SIMILARITY.
FT DISULFID 2478 2493 BY SIMILARITY.
FT DISULFID 2495 2504 BY SIMILARITY.
FT DISULFID 2511 2522 BY SIMILARITY.
FT DISULFID 2539 2631 BY SIMILARITY.
FT DISULFID 2607 2623 BY SIMILARITY.
FT DISULFID 2638 2681 BY SIMILARITY.
FT DISULFID 2667 2694 BY SIMILARITY.
FT CARBOHYD 57 57 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 330 330 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 692 692 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 758 758 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 802 802 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 805 805 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1257 1257 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1435 1435 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1633 1633 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1660 1660 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1684 1684 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1738 1738 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1848 1848 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2004 2004 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2409 2409 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2711 2711 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2721 2721 N-LINKED (GLCNAC. .) (POTENTIAL).
FT VARSPPLIC 349 2431 Missing (in isoform V3).
FT VARSPPLIC 2697 2738 /FTId=VSP_003091.
FT PSAYQRTYSKYLKNSSSVKDINSINTSKHEHRSRRWOETR
FT R -> RKWSFRKNGQPCFNKY (in isoform Vint).
FT /FTId=VSP_003092.
FT AEREC -> NSARG (IN REF. 4).
FT CONFLICT 2535 2539
FT SEQUENCE 2738 AA; 300004 MW; 12CA626D58BD8C6A CRC64;
SQ
Query Match 15.6%; Score 174; DB 1; Length 2738;
Best Local Similarity 28.5%; Pred. No. 3.9e-07;
Matches 47; Conservative 23; Mismatches 55; Indels 40; Gaps 8;
QY 17 QRPCYKVIYFHDTSRRLNFEAEACRRDGGQLVSISEDEQKLEKFIENLLPSDGF- 75
Db 2519 QGQCYK--YF---AHRRTWDAARECRLQGAHLTSILSHEEQMFVNRV-----GHDIQ 2566
QY 76 WIGLRRREKQSNSTACQDLYAWTDGSIQFRNHYVDEP-----SCGSEVCVMVHQPSAP 131
Db 2567 WIGL-----NDKMFEDHFRWTDGSALEYENWRPNQDPDSFFSAGEDCVVIWHENG-- 2616
QY 132 AGIGGPPYMFQWDDRCNMKNFICKYS-----DEKPAVPSREAGE 172
Db 2617 -----QWNVPCNYHLTYTCKKGTGTVACGQPPVVENAKTFGK 2652

RESULT 9
PGCV_MOUSE STANDARD; PRT; 3358 AA.
AC Q62059; Q62058; Q9CUU0;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Versican core protein precursor (Large fibroblast proteoglycan)
DE (Chondroitin sulfate proteoglycan core protein 2) (PG-M).
GN CSPG2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A. (ISOFORMS V0; V1 AND V2).
RC STRAIN=C57BL/6, and Swiss Webster; TISSUE=Brain;
RX MEDLINE=95122551; PubMed=7822336;
RA Ito K., Shinomura T., Zako M., Ujita M., Kimata K.;
RT "Multiple forms of mouse PG-M, a large chondroitin sulfate
RT proteoglycan generated by alternative splicing.";
RL J. Biol. Chem. 270:958-965(1995).
RN [2]
SEQUENCE FROM N.A. (ISOFORM V3).
RP STRAIN=C57BL/6;
RC STRAIN=C57BL/6;
RX MEDLINE=95181355; PubMed=7876137;
RA Zako M., Shinomura T., Ujita M., Ito K., Kimata K.;
RT "Expression of PG-M(V3), an alternatively spliced form of PG-M
RT without a chondroitin sulfate attachment in region in mouse and human
RT tissues.";
RL J. Biol. Chem. 270:3914-3918(1995).
RN [3]
SEQUENCE OF 1-1692 FROM N.A. (ISOFORM V1).
RC STRAIN=C57BL/6J; TISSUE=Skin;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [4]
INTERACTION WITH FBLN1.
RX PubMed=10400671;
RA Asperberg A., Adam S., Kostka G., Timpl R., Heinegaard D.;
RT "Fibulin-1 is a ligand for the C-type lectin domains of aggrecan and
RT versican.";
RL J. Biol. Chem. 274:20444-20449(1999).
CC -!- FUNCTION: May play a role in intercellular signaling and in
CC connecting cells with the extracellular matrix. May take part in
CC the regulation of cell motility, growth and differentiation. Binds
CC hyaluronic acid.
CC -!- SUBUNIT: Interacts with FBLN1.
CC -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=4;
CC Comment=Additional isoforms seem to exist;
CC Name=V0;
CC IsoId=Q62059-1; Sequence=Displayed;

CC Name=V1;
CC IsoId=Q62059-2; Sequence=VSP_003087, VSP_003088;
CC Name=V2;
CC IsoId=Q62059-3; Sequence=VSP_003089;
CC Name=V3;
CC IsoId=Q62059-4; Sequence=VSP_003087, VSP_003090;
CC TISSUE SPECIFICITY: V2 is found only in brain.
CC -!- DEVELOPMENTAL STAGE: Disappears after the cartilage development.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -!- SIMILARITY: Contains 2 link domains.
CC -!- SIMILARITY: Contains 2 EGF-like domains.
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC -!- SIMILARITY: Contains 1 Sushi (SCR) domain.
CC -!- SIMILARITY: BELONGS TO THE AGGRECAN/VERSICAN PROTEOGLYCAN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D16263; BAA03796.1; --
DR EMBL; D28599; --; NOT ANNOTATED_CDS.
DR EMBL; D32040; BAA06802.1; --
DR EMBL; AK014525; BAB29411.1; --
DR HSSP; P01132; IEPG.
DR MGD; MGI:102889; Cspg2.
DR InterPro; IPR000152; A6x hydroxyl.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR007110; Ig_Like.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR000538; Link.
DR InterPro; IPR000436; Sushi_SCR_CCP.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00059; lectin_c; 1.
DR Pfam; PF00084; sushi; 1.
DR Pfam; PF00193; Xlink; 2.
DR PRINTS; PR01265; LINKMODULE.
DR ProDom; PD000918; Link; 2.
DR SMART; SM00032; CCP; 1.
DR SMART; SM00034; CLECT; 1.
DR SMART; SM00179; EGF_CA; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00445; LINK; 2.
DR PROSITE; PS00010; ASX HYDROXYL; 1.
DR PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01187; EGF_CA; 1.
DR PROSITE; PS50835; IG_Like; 1.
DR PROSITE; PS01241; LINK; 2.
KW Glycoprotein; Proteoglycan; Lectin; Extracellular matrix; Sushi;
KW Signal; Repeat; EGF-like domain; Calcium; Immunoglobulin domain;
KW Hyaluronic acid; Alternative splicing.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 3358 VERSICAN CORE PROTEIN.
FT DOMAIN 21 146 IG-LIKE V-TYPE.
FT DOMAIN 167 244 LINK 1.
FT DOMAIN 265 346 LINK 2.
FT DOMAIN 348 1308 GAG-ALPHA
(GLUCOSAMINOGLYCAN ATTACHMENT DOMAIN).
FT DOMAIN 1309 3052 GAG-BETA.
FT DOMAIN 3052 3088 EGF-LIKE 1.
FT DOMAIN 3090 3126 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 3139 3253 C-TYPE LECTIN.

Db 3390 WIGL-----NDKMFERDFRTDGSPLQYENWRPNQDPSFFSAGEDCVIIWHENG-- 3439

QY 132 AGIGGPFYMFQWDDRCNMKNFNICKYS-----DEKPAVPSREAGE 172

Db 3440 -----QWNVPCNYHLTYTCKKGTVACGQPPVVENAKTFGK 3475

RESULT 12

CD93_RAT

ID CD93_RAT STANDARD; PRT; 643 AA.

AC Q9ET61; Q9JIZ6;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Complement component Clq receptor precursor (Complement component 1, q subcomponent, receptor 1) (ClqRp) (ClqR(p)) (Clq/MBL/SPA receptor)

DE subcomponent, receptor 1) (ClqRp) (ClqR(p)) (Clq/MBL/SPA receptor)

DE CD93 antigen) (Cell surface antigen AA4).

GN ClQR1 OR CD93 OR ClQRP.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=PVG; TISSUE=Natural killer cells;

RX MEDLINE=20545218; PubMed=11093152;

RA Dean Y.D., McGreal E.P., Akatsu H., Gasque P.;

RT "Molecular and cellular properties of the rat AA4 antigen, a C-type lectin-like receptor with structural homology to thrombomodulin.";

RL J. Biol. Chem. 275:34382-34392(2000).

CC -!- FUNCTION: Receptor (or element of a larger receptor complex) for Clq, mannose-binding lectin (MBL2) and pulmonary surfactant protein A (SPA). May mediate the enhancement of phagocytosis in monocytes and macrophages upon interaction with soluble defense collagens. May play a role in intercellular adhesion.

CC -!- SUBCELLULAR LOCATION: Type I membrane protein.

CC -!- TISSUE SPECIFICITY: Widely expressed. Highly expressed in lung and heart. Expressed at lower level in brain, thymus, liver, spleen, intestine, kidney, adrenal gland, muscle and testis. Expressed on endothelial cells, platelets, undifferentiated monocytes and circulating natural killer cells.

CC -!- PTM: N- and O-glycosylated (By similarity).

CC -!- SIMILARITY: Contains 1 C-type lectin family domain.

CC -!- SIMILARITY: Contains 5 EGF-like domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; AF136537; AAC01572.1; -.

DR EMBL; AF160978; AAF80402.1; -.

DR HSSP; P35555; 1EMN.

DR GO; GO:0016021; C:integral to membrane; ISS.

DR GO; GO:0004872; F:receptor activity; ISS.

DR GO; GO:0016337; P:cell-cell adhesion; ISS.

DR GO; GO:0042116; P:macrophage activation; ISS.

DR GO; GO:0006909; P:phagocytosis; ISS.

DR InterPro; IPR000152; Asx hydroxyl.

DR InterPro; IPR001881; EGF_Ca.

DR InterPro; IPR006209; EGF_like.

DR InterPro; IPR001304; Lectin_C.

DR Pfam; PF00008; EGF; 5.

DR Pfam; PF00059; lectin_c; 1.

DR SMART; SM00034; CLECT; 1.

DR SMART; SM00179; EGF_CA; 3.

DR PROSITE; PS00010; ASX_HYDROXYL; 3.

DR PROSITE; PS00615; C_TYPE_LECTIN_1; FALSE_NEG.

DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.

DR PROSITE; PS01186; EGF_2; 3.

DR PROSITE; PS01187; EGF_CA; 3.

KW Cell adhesion; Receptor; Repeat; Signal; Transmembrane;

KW EGF-like domain; Lectin; Glycoprotein.

FT SIGNAL 1 23 POTENTIAL.

FT CHAIN 24 643 COMPLEMENT COMPONENT C1Q RECEPTOR.

FT DOMAIN 24 571 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 572 592 POTENTIAL.

FT DOMAIN 593 643 CYTOPLASMIC (POTENTIAL).

FT DOMAIN 31 173 C-TYPE LECTIN.

FT DOMAIN 257 298 EGF-LIKE 1.

FT DOMAIN 299 341 EGF-LIKE 2.

FT DOMAIN 342 381 EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).

FT DOMAIN 382 423 EGF-LIKE 4, CALCIUM-BINDING (POTENTIAL).

FT DOMAIN 424 462 EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).

FT DISULFID 261 272 BY SIMILARITY.

FT DISULFID 268 282 BY SIMILARITY.

FT DISULFID 284 297 BY SIMILARITY.

FT DISULFID 303 314 BY SIMILARITY.

FT DISULFID 308 325 BY SIMILARITY.

FT DISULFID 327 340 BY SIMILARITY.

FT DISULFID 346 355 BY SIMILARITY.

FT DISULFID 351 364 BY SIMILARITY.

FT DISULFID 366 380 BY SIMILARITY.

FT DISULFID 386 397 BY SIMILARITY.

FT DISULFID 393 406 BY SIMILARITY.

FT DISULFID 408 422 BY SIMILARITY.

FT DISULFID 428 437 BY SIMILARITY.

FT DISULFID 433 446 BY SIMILARITY.

FT DISULFID 448 461 BY SIMILARITY.

FT CARBOHYD 322 322 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 498 498 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CONFLICT 417 417 E -> K (IN REF. 2).

SQ SEQUENCE 643 AA; 68781 MW; 9AE4C933AD943DB6 CRC64;

Query Match 14.8%; Score 165; DB 1; Length 643;

Best Local Similarity 25.2%; Pred. No. 4.2e-07;

Matches 52; Conservative 35; Mismatches 73; Indels 46; Gaps 10;

QY 3 GRLLSGQP-----VCRGGTQPCYKVIYFHDTSRRLNFEAEAKRRDGGQLVSI ESD 56

Db 15 GQLWAGAAADSEAVVCEG---TACYTAHW-----GKLSAAEAQHRCNENGNLATVKSEE 66

QY 57 EQKLI EKFIENLL---PSD---GDFWIGLRREEKQSNSTACQDLYAWT-DGSISQFRN 108

Db 67 EARHVQEALAQLLKTKAPSETKIGKFWIGLQREKCKTYHDLPMKGFWSVGGGEDTYSN 126

QY 109 WY-VDEPSCGSEVCVVMY-----HQPSAPAGIGGPPYMFQWDDRC-----NMKNVF 153

Db 127 WYKASKSSCISKRCVSLILDLSLKPHPSHLP-----KWHESPCGTPDAPGNSIEGF 177

QY 154 IKYSDKPAVPSREAGEETELTTP 179

Db 178 LCKFNFKGMCSPALALGGPGQLTYTTP 203

RESULT 13

PGCA_BOVIN

ID PGCA_BOVIN STANDARD; PRT; 2364 AA.

AC P13608; P79117; Q28159;

DT 01-JAN-1990 (Rel. 13, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE AggreCAN core protein precursor (Cartilage-specific proteoglycan core protein) (CSPCP).

FT DISULFID 297 318 BY SIMILARITY.
FT DISULFID 509 580 BY SIMILARITY.
FT DISULFID 533 554 BY SIMILARITY.
FT DISULFID 607 682 BY SIMILARITY.
FT DISULFID 631 652 BY SIMILARITY.
FT DISULFID 2117 2128 BY SIMILARITY.
FT DISULFID 2182 2274 BY SIMILARITY.
FT DISULFID 2250 2266 BY SIMILARITY.
FT DISULFID 2281 2324 BY SIMILARITY.
FT DISULFID 2310 2337 BY SIMILARITY.
FT CARBOHYD 126 126 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 239 239 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 333 333 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 387 387 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 611 611 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 667 667 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 2114 2150 Missing (in isoform 2).
FT /FTID=VSP 003072.
SQ SEQUENCE 2364 AA; 246359 MW; 6FF83763420C3D4C CRC64;
Query Match 14.2%; Score 158.5; DB 1; Length 2364;
Best Local Similarity 26.8%; Pred. No. 7.8e-06;
Matches 48; Conservative 29; Mismatches 59; Indels 43; Gaps 10;
QY 9 QPVCRGG---TORPCYKVIYFHDTSRRLLNFEEAKEACRRDGGQLVSEIEDEQKLEKFI 65
Db 2151 QKLCIEGWTQFGHCYR--HFPD---RATWVDAESQCRKQKSHLSSIVTPEEQ----EFV 2201
QY 66 ENLLPSDGDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNWWYVDEP----SGSEV 120
Db 2202 NN---NAQDYQWIGL-----NDKTIEGDFRWSGHSLSQFENWRPNQPNFFATGEDC 2250
QY 121 CVVMYHQPSAPAGIGGPPYFQWNDRCNMKNFICKYS----DEKPAVPSREAEGETE 175
Db 2251 VVMIWHEKG-----EWNQVPCNYQLPFTCKKGTACGEPVVEHARIFGQKQD 2298
RESULT 14
PGCA_HUMAN STANDARD; PRT; 2415 AA.
AC P16112; Q13650;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Aggrecan core protein precursor (Cartilage-specific proteoglycan core protein)
DE protein) (CSPCP) (Chondroitin sulfate proteoglycan core protein 1).
GN AGC1 OR CSPG1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 3).
RC TISSUE=Chondrocytes;
RX MEDLINE=91093289; PubMed=1985970;
RA Doege K.J., Sasaki M., Kimura T., Yamada Y.;
RT "Complete coding sequence and deduced primary structure of the human
RT cartilage large aggregating proteoglycan, aggrecan. Human-specific
RT repeats, and additional alternatively spliced forms.";
RL J. Biol. Chem. 266:894-902(1991).
RN [2]
RP SEQUENCE OF 1778-2415 FROM N.A. (ISOFORM 2).
RC TISSUE=Chondrocytes;
RA Dudhia J., Hardingham T.E.;
RL Submitted (JAN-1990) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 1936-2415 FROM N.A. (ISOFORM 1).
RX MEDLINE=89380154; PubMed=2789216;
RA Baldwin C.T., Reginato A.M., Prockop D.J.;
RT "A new epidermal growth factor-like domain in the human core protein
RT for the large cartilage-specific proteoglycan. Evidence for
RT alternative splicing of the domain.";
RL J. Biol. Chem. 264:15747-15750(1989).

RN [4]
RP SEQUENCE OF 764-864 FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=95128522; PubMed=7827755;
RA Barry F.P., Neame P.J., Sasse J., Pearson D.;
RT "Length variation in the keratan sulfate domain of mammalian
RT aggrecan.";
RL Matrix Biol. 14:323-328(1994).
CC -1- FUNCTION: THIS PROTEOGLYCAN IS A MAJOR COMPONENT OF EXTRACELLULAR
CC MATRIX OF CARTILAGINOUS TISSUES. A MAJOR FUNCTION OF THIS PROTEIN
CC IS TO RESIST COMPRESSION IN CARTILAGE. IT BINDS AVIDLY TO
CC HYALURONIC ACID VIA AN AMINO-TERMINAL GLOBULAR REGION.
CC -1- SUBUNIT: Interacts with FBLN1 (By similarity).
CC -1- SUBCELLULAR LOCATION: Secreted; extracellular matrix (By
CC similarity).
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Comment=Additional isoforms seem to exist;
CC Name=1;
CC IsoId=P16112-1; Sequence=Displayed;
CC Name=2;
CC IsoId=P16112-2; Sequence=VSP_003074;
CC Name=3;
CC IsoId=P16112-3; Sequence=VSP_003074, VSP_003075;
CC DOMAIN: TWO GLOBULAR DOMAINS, G1 AND G2, COMPRISE THE AMINO
CC TERMINUS OF THE PROTEOGLYCAN, WHILE ANOTHER GLOBULAR REGION, G3,
CC MAKES UP THE COOH TERMINUS. G1 CONTAINS LINK DOMAINS AND THUS
CC CONSISTS OF THREE DISULFIDE-BONDED LOOP STRUCTURES DESIGNATED AS
CC THE A, B, B' MOTIFS. G2 IS SIMILAR TO G1. THE KERATAN SULFATE (KS)
CC AND THE CHONDROITIN SULFATE (CS) ATTACHMENT DOMAINS LIE BETWEEN G2
CC AND G3.
CC -1- PTM: CONTAINS MOSTLY CHONDROITIN SULFATE, BUT ALSO KERATAN SULFATE
CC CHAINS, N-LINKED AND O-LINKED OLIGOSACCHARIDES.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -1- SIMILARITY: Contains 4 link domains.
CC -1- SIMILARITY: Contains 1 EGF-like domain.
CC -1- SIMILARITY: Contains 1 C-type lectin family domain.
CC -1- SIMILARITY: Contains 1 Sushi (SCR) domain.
CC -1- SIMILARITY: BELONGS TO THE AGGRECAN/VERSICAN PROTEOGLYCAN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; M55172; AAA62824.1; -.
DR EMBL; J05062; AAA35726.1; -.
DR EMBL; X17406; CAA35463.1; -.
DR EMBL; S74659; AAC60643.2; -.
DR PIR; A39086; A39086.
DR HSSP; P98066; 1TSG.
DR Genew; HGNC:319; AGC1.
DR MIM; 155760; -.
DR GO; GO:0005204; F:chondroitin sulfate proteoglycan; TAS.
DR InterPro; IPR002353; Antifreeze1.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003006; Ig MHC.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR000538; Link.
DR InterPro; IPR003324; SGXXSG.
DR InterPro; IPR000436; Sushi_SCR_CCP.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF00059; lectin c; 1.
DR Pfam; PF02339; SGXXSG; 71.
DR Pfam; PF00084; sushi; 1.
DR Pfam; PF00193; Xlink; 4.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 22, 2003, 16:06:34 ; Search time 10.0151 Seconds
(without alignments)
1978.090 Million cell updates/sec

Title: US-09-887-855-2_COPY_22_227
Perfect score: 1115
Sequence: 1 ATGRLLSGQPVCRGGTQRPC.....EEDAKKTFKESREAAALNLAY 206

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	182	16.3	1456	1 A36563	mannose receptor p
2	178.5	16.0	1455	1 A48925	mannose receptor p
3	177	15.9	1268	2 S52781	neurocan - mouse
4	174.5	15.7	1643	2 T14274	versican precursor
5	174.5	15.7	3381	2 T42389	versican precursor
6	174	15.6	1257	2 S28764	neurocan precursor
7	174	15.6	2397	1 A55535	versican precursor
8	174	15.6	2409	1 A60979	versican precursor
9	171	15.3	3562	2 A47171	chondroitin sulfat
10	170.5	15.3	1479	2 T42710	mannose receptor,
11	158.5	14.2	1340	2 A39808	proteoglycan core
12	158.5	14.2	2327	2 T42630	aggrecan - bovine
13	158.5	14.2	2415	1 A39086	aggrecan precursor
14	154.5	13.9	612	2 B42755	E-selectin precurs
15	153.5	13.8	2124	2 A28452	proteoglycan core
16	152	13.6	912	2 A54423	brevican precursor
17	149.5	13.4	459	2 T24425	hypothetical prote
18	149	13.4	321	1 LNHUER	IgE Fc receptor II
19	148.5	13.3	330	2 T46256	brevican - human (
20	148.5	13.3	2132	1 A55182	aggrecan precursor
21	148	13.3	253	2 E89130	protein F52E1.2 [i
22	147	13.2	883	2 S57653	brevican precursor
23	146.5	13.1	162	1 LNRCl	lectin BRA3-1 prec
24	146	13.1	2109	1 I50421	aggrecan precursor
25	145.5	13.0	742	2 JC7595	scavenger receptor
26	145	13.0	883	2 S49126	brevican precursor
27	144.5	13.0	173	2 S10548	lectin - barnacle
28	144.5	13.0	372	2 S23936	L-selectin precurs
29	144.5	13.0	404	2 A46274	HIV gp120-binding

30	143.5	12.9	129	2	JC4329	coagulation factor
31	143.5	12.9	372	1	A32375	L-selectin precurs
32	143.5	12.9	463	2	T26655	hypothetical prote
33	142.5	12.8	131	2	JC5058	bitiscetin alpha c
34	142	12.7	1487	2	S48719	phospholipase-A(2)
35	141.5	12.7	331	1	LNMSER	IgE Fc receptor, 1
36	140.5	12.6	162	1	LNRC3	lectin BRA3-2 prec
37	140	12.6	370	2	S22124	L-selectin precurs
38	139	12.5	248	1	LNHUPS	pulmonary surfacta
39	139	12.5	248	1	LNHUP6	pulmonary surfacta
40	139	12.5	248	1	LNHUP1	pulmonary surfacta
41	139	12.5	283	1	LNFHLS	lectin precursor -
42	138.5	12.4	152	2	JC4690	coagulation factor
43	138.5	12.4	202	2	JC4031	tetranectin precur
44	138	12.4	280	2	T29200	hypothetical prote
45	137.5	12.3	309	1	S34198	IgE Fc receptor II

ALIGNMENTS

RESULT 1

A36563
mannose receptor precursor - human
C;Species: Homo sapiens (man)
C;Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text change 10-Sep-1999
C;Accession: A36563; A60926; A44255; B44255; C44255; D44255; E44255; F44255; G44255; H:
R;Taylor, M.E.; Conary, J.T.; Lennartz, M.R.; Stahl, P.D.; Drickamer, K.
J. Biol. Chem. 265, 12156-12162, 1990
A;Title: Primary structure of the mannose receptor contains multiple motifs resembling
A;Reference number: A36563; MUID:90324192; PMID:2373685
A;Accession: A36563
A;Molecule type: mRNA
A;Residues: 1-1456 <TAY>
A;Cross-references: GB:J05550; NID:g188675; PIDN:AAAS9868.1; PID:g188676
A;Note: parts of this sequence, including the amino end of the mature protein, were CO
R;Ezekowitz, R.A.B.; Sastry, K.; Bailly, P.; Warner, A.
J. Exp. Med. 172, 1785-1794, 1990
A;Title: Molecular characterization of the human macrophage mannose receptor: demonstr
A;Reference number: A60926; MUID:91079783; PMID:2258707
A;Accession: A60926
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-1333,'T',1335-1456 <EZE>
A;Cross-references: GB:X55635
A;Note: translation of the nucleotide sequence is incomplete
A;Note: in the authors' translation additional residues Pro-Glu-Ile are shown after 49
R;Kim, S.J.; Ruiz, N.; Bezouska, K.; Drickamer, K.
Genomics 14, 721-727, 1992
A;Title: Organization of the gene encoding the human macrophage mannose receptor (MRC1
A;Reference number: A44255; MUID:93052405; PMID:1294118
A;Accession: A44255
A;Status: preliminary; nucleic acid sequence not shown; not compared with conceptual t
A;Molecule type: DNA
A;Residues: 155-233,'KSAL',238-283;346-428;492-569;631-714,716-719;783-820,'N',822-865
A;Note: sequence extracted from NCBI backbone (NCBIP:118415, NCBIP:118421, NCBIP:11842
C;Genetics:
A;Gene: GDB:MRC1
A;Cross-references: GDB:133759; OMIM:153618
A;Map position: 10p13-10p13
C;Superfamily: phospholipase A2 receptor; C-type lectin homology; fibronectin type II :
C;Keywords: duplication; lectin; tandem repeat; transmembrane protein
F;1-18/Domain: signal sequence #status predicted <SIG>
F;168-209/Domain: fibronectin type II repeat homology <2F1>
F;223-340/Domain: C-type lectin homology <LCH1>
F;362-486/Domain: C-type lectin homology <LCH2>
F;945-1079/Domain: C-type lectin homology <LCH3>

Query Match 16.3%; Score 182; DB 1; Length 1456;
Best Local Similarity 25.5%; Pred. No. 9.1e-08;
Matches 52; Conservative 37; Mismatches 61; Indels 54; Gaps 8;

QY 21 YKVIYFHDTSRLNFEEAKEACRRDGGQLVSISEDEQKLIKFIENLLPSDGFWIGLR 80

Db 807 YKDYQYFSKEKETMDNARAFCKRNFGLVSIQSESEKKFLWKYV-NRNDASAYFIGLL 865
QY 81 RREEKQNSTACQDLYAWTDGSIISQFRNMYVDPSGCS--EVCVVMYHQPSAPAGIGPY 138
Db 866 ISLDKK-----FAWMDGSKVDYVSWATGEPNFANEDENCVTMY-----SNSGF---- 908
QY 139 MFQWNDRCNMKNPFICKYSDEK----PAVPSREAEGETELTTPVLPEETQE----- 187
Db 909 ---WNDINCGYPNAFICQRHNSSINATTVM-----TMPSPVSGCKEGWNFYSN 954
QY 188 -----EDAKKTFKESREAAAL 202
Db 955 KCFKIFGFMEERKNWQEARACI 978
RESULT 2
A48925
mannose receptor precursor, macrophage - mouse
C;Species: Mus musculus (house mouse)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A48925; S21320; PC2245
R;Harris, N.; Super, M.; Rits, M.; Chang, G.; Ezekowitz, R.A.
Blood 80, 2363-2373, 1992
A;Title: Characterization of the murine macrophage mannose receptor: demonstration that
A;Reference number: A48925; MUID:93043353; PMID:1421407
A;Accession: A48925
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-1455 <HAR>
A;Experimental source: peritoneal macrophage
A;Note: sequence extracted from NCBI backbone (NCBIP:118733)
R;Harris, N.; Super, M.; Rits, M.; Chang, G.; Ezekowitz, R.B.
submitted to the EMBL Data Library, April 1992
A;Description: Characterization of the murine macrophage mannose receptor: Demonstration
on.
A;Reference number: S21320
A;Accession: S21320
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-302,'W',303-1117,'E',1119-1455 <HA2>
A;Cross-references: EMBL:211974; NID:G52997; PIDN:CAA78028.1; PID:G52998
R;Harris, N.; Peters, L.L.; Eicher, E.M.; Rits, M.; Rasperry, D.; Eichbaum, Q.G.; Super
Biochem. Biophys. Res. Commun. 198, 682-692, 1994
A;Title: The exon-intron structure and chromosomal localization of the mouse macrophage
A;Reference number: PC2245; MUID:94128116; PMID:8297379
A;Accession: PC2245
A;Molecule type: mRNA
A;Residues: 35-105 <HA3>
C;Genetics:
A;Gene: Mrc1
A;Map position: 2
C;Superfamily: phospholipase A2 receptor; C-type lectin homology; fibronectin type II re
C;Keywords: membrane protein; receptor
F;168-209/Domain: fibronectin type II repeat homology <2F9>
F;361-485/Domain: C-type lectin homology <LCH1>
F;943-1077/Domain: C-type lectin homology <LCH2>
Query Match 16.0%; Score 178.5; DB 1; Length 1455;
Best Local Similarity 25.9%; Pred. No. 1.8e-07;
Matches 53; Conservative 33; Mismatches 66; Indels 53; Gaps 8;
QY 21 YKVIYFHDTSRRLNFEFEAKEACRRDGGQLVSIQSESEKFLWKYI-NKNGGQSPYFIGML 80
Db 806 YKDYQYFSKEKETMDNARRFCKKNFGDLATIKSESEKKFLWKYI-NKNGGQSPYFIGML 864
QY 81 RREEKQNSTACQDLYAWTDGSIISQFRNMYVDPSGCS--EVCVVMYHQPSAPAGIGPY 138
Db 865 ISMDKK-----FIWMDGSKVDVFAWATGEPNFANDDENCVTMY-----TNSGF---- 907
QY 139 MFQWNDRCNMKNPFICK----YSDEKPAVPSREAEGETELTTPVLPEETQE----- 187
Db 908 ---WNDINCGYPNPFICQRHNSSINATAMP-----TTPTPGCKEGWHLKYK 953

QY 188 -----EDAKKTFKESREAAALNL 204
Db 954 CFKIFGFANBEKKSWQDARQACKGL 978
RESULT 3
S52781
neurocan - mouse
C;Species: Mus musculus (house mouse)
C;Date: 19-May-1995 #sequence_revision 21-Jul-1995 #text_change 04-Feb-2000
C;Accession: S52781
R;Rauch, U.; Forsberg, N.; Kulbe, G.; Arnold-Ammer, I.; Faessler, R.
submitted to the EMBL Data Library, February 1995
A;Description: Amino acid sequence of mouse neurocan and their different e
A;Reference number: S52781
A;Accession: S52781
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1268 <RAU>
A;Cross-references: EMBL:X84727; NID:G758629; PIDN:CAA59216.1; PID:G758630
C;Superfamily: aggrecan; C-type lectin homology; complement factor H repeat homology; E
F;176-253/Domain: link protein repeat homology <LNK1>
F;274-355/Domain: link protein repeat homology <LNK2>
F;964-995/Domain: EGF homology <EGF>
F;1040-1160/Domain: C-type lectin homology <LCH>
F;1167-1223/Domain: complement factor H repeat homology <PHD>
Query Match 15.9%; Score 177; DB 2; Length 1268;
Best Local Similarity 31.5%; Pred. No. 2.1e-07;
Matches 45; Conservative 17; Mismatches 49; Indels 32; Gaps 6;
QY 17 QRPCYKVIYFHDTSRRLNFEFEAKEACRRDGGQLVSIQSESEKFLWKYIENLLPSDGDGF 76
Db 1048 QGHCYR--YF---AHRRAWEDAERDCRRRAGHLTSVHSPEEHKFINSF-----GHENSW 1096
QY 77 IGLRRREKQNSTACQDLYAWTDGSIISQFRNMYVDEPS---CGSEVCVVMYHQPSAPAG 133
Db 1097 IGLNDRTVERD-----FQWTDNTGLQYENWREKQPDNFFAGGEDCVVMVAHESG--- 1145
QY 134 IGGPYMFQWNDRCNMKNPFICK 156
Db 1146 -----RWNDVPCYNLPYVCK 1161

RESULT 4
T14274
versican precursor, splice form V2 - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 05-May-2000
C;Accession: T14274
R;Schmalfeldt, M.; Dours-Zimmermann, M.T.; Winterhalter, K.H.; Zimmermann, D.R.
J. Biol. Chem. 273, 15758-15764, 1998
A;Title: Versican V2 is a major extracellular matrix component of the mature bovine bra
A;Reference number: Z17954; MUID:98288320; PMID:9624174
A;Accession: T14274
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1643 <SCH>
A;Cross-references: EMBL:AF060458; NID:G3253303; PID:G3253304; PIDN:AAC24360.1
A;Experimental source: brain
C;Keywords: glycoprotein
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-1643/Product: versican, splice form V2 #status predicted <MAT>
F;57,331,352,817,965,1017,1333,1616,1626/Binding site: carbohydrate (Asn) (covalent) #s
Query Match 15.7%; Score 174.5; DB 2; Length 1643;
Best Local Similarity 25.5%; Pred. No. 4.8e-07;
Matches 59; Conservative 26; Mismatches 71; Indels 75; Gaps 11;
QY 17 QRPCYKVIYFHDTSRRLNFEFEAKEACRRDGGQLVSIQSESEKFLWKYIENLLPSDGDGF- 75
Db 1424 QGCQYK--YF---AHRRTWDAARECRLOGAHLTSLSHEEQMFVNRV-----GHDYQ 1471

C;Accession: T42710
R;Wu, K.; Yuan, J.; Lasky, L.A.
J. Biol. Chem. 271, 21323-21330, 1996
A;Title: Characterization of a novel member of the macrophage mannose receptor type C 1e
A;Reference number: 222235; MUID:96355501; PMID:8702911
A;Accession: T42710
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1479 <WUK>
A;Cross-references: EMBL:U56734; NID:g1336073; PID:g1336074; PIDN:AAC52729.1
C;Superfamily: phospholipase A2 receptor; C-type lectin homology; fibronectin type II re
C;Keywords: membrane protein; receptor
F;186-227/Domain: fibronectin type II repeat homology <2FR>

Query Match 15.3%; Score 170.5; DB 2; Length 1479;
Best Local Similarity 31.4%; Pred. No. 9.6e-07;
Matches 48; Conservative 22; Mismatches 48; Indels 35; Gaps 7;

QY 7 SGQVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIIESEDEQKLIKFI 66
Db 384 SWQFF-----QGHCYRL-----QA EKRSWQESKRACLRGGDLLSIHSMAELEFITKQIK 433

QY 67 NLLPSDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNMYVDEPS---CGSEVCVV 123
Db 434 QEVE---ELWIGL-----NDLKLQMNFEWSGSLVSFTHWHFEPNFRDLSLEDCVT 482

QY 124 MYHQSAPAGIGGPFYMFQWNDRCNMKNPFICK 156
Db 483 IW-----GPEG-----RWNDSPCNQSLPSICK 504

RESULT 11
A39808
proteoglycan core protein, cartilage - bovine (fragments)
N;Alternate names: aggrecan; aggregating cartilage proteoglycan
C;Species: Bos primigenius taurus (cattle)
C;Date: 20-Mar-1992 #sequence_revision 23-Mar-1995 #text_change 13-Aug-1999
C;Accession: A34234; A27752; A39808; A27751; E29164; B27751; D27751; E27751; F27
R;Antonsson, P.; Heinegard, D.; Oldberg, A.
J. Biol. Chem. 264, 16170-16173, 1989
A;Title: The keratan sulfate-enriched region of bovine cartilage proteoglycan consists c
A;Reference number: A34234; MUID:89380219; PMID:2528543
A;Accession: A34234
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 128-621 <ANT>
A;Cross-references: GB:J05028
R;Oldberg, A.; Antonsson, P.; Heinegard, D.
Biochem. J. 243, 255-259, 1987
A;Title: The partial amino acid sequence of bovine cartilage proteoglycan,deduced from a
A;Reference number: A27752; MUID:87270630; PMID:3111460
A;Accession: A27752
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 622-1340 <OLD>
R;Sandy, J.D.; Boynton, R.E.; Flannery, C.R.
J. Biol. Chem. 266, 8198-8205, 1991
A;Title: Analysis of the catabolism of aggrecan in cartilage explants by quantitation of
A;Reference number: A39808; MUID:91217051; PMID:2022637
A;Accession: A39808
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-28;59-82;131-137,'QSET',142-149;196-207;226-249;1137-1143;1252-1267;1274-1
R;Perin, J.P.; Bonnet, F.; Jolles, P.
FEBS Lett. 206, 73-77, 1986
A;Title: Structural relationship between link proteins and proteoglycan monomers.
A;Reference number: A27751; MUID:87005253; PMID:3530809
A;Accession: A27751
A;Molecule type: protein
A;Residues: 29-58;74-130;174-175,'A',177-204;208-225 <PER>
R;Perin, J.P.; Bonnet, F.; Jolles, J.; Jolles, P.
FEBS Lett. 176, 37-42, 1984
A;Title: Sequence data concerning the protein core of the cartilage proteoglycan monomer

A;Reference number: A91327; MUID:85027710; PMID:6489519
A;Accession: E29164
A;Molecule type: protein
A;Residues: 1230-1249 <PE2>
C;Superfamily: aggrecan; C-type lectin homology; complement factor H repeat homology; I
C;Keywords: glycoprotein
F;8-28/Domain: link protein repeat homology (fragment) <LNK1>
F;29-58/Domain: link protein repeat homology (fragment) <LNK2>
F;80-146/Domain: link protein repeat homology (fragments) <LNK3>
F;167-248/Domain: link protein repeat homology <LNK4>
F;1130-1250/Domain: C-type lectin homology <LCH>
F;1257-1313/Domain: complement factor H repeat homology <FHD>

Query Match 14.2%; Score 158.5; DB 2; Length 1340;
Best Local Similarity 26.8%; Pred. No. 9.9e-06;
Matches 48; Conservative 29; Mismatches 59; Indels 43; Gaps 10;

QY 9 QPVCRRGG--TORPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIIESEDEQKLIKFI 65
Db 1127 QKLCEEGWTKFQGHCHYR--HFPD---RATWVDAESQCRKQOSHLSIVTPEEQ----EFV 1177

QY 66 ENLLPSDGF-WIGLRRREEKQSNSTACQDLYAWTDGSIQFRNMYVDEP-----SCGSEV 120
Db 1178 NN---NAQDYQWIGL-----NDKTIEGDFRWSDGHSLOFENWRPNQDNFFATGEDC 1226

QY 121 CVVMYHQSAPAGIGGPFYMFQWNDRCNMKNPFICKYS-----DEKPAVPSREAEGETE 175
Db 1227 VVMIWHEKG-----EWNDVPCNYQLPFTCKKGTVACGEPVVEHARIFGQKKD 1274

RESULT 12
T42630
aggrecan - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000
C;Accession: T42630
R;Hering, T.M.; Kollar, J.; Huynh, T.D.
submitted to the EMBL Data Library, September 1996
A;Description: Complete coding sequence of bovine aggrecan: comparative structural ana.
A;Reference number: Z22182
A;Accession: T42630
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-2327 <HER>
A;Cross-references: EMBL:U76615; NID:g1730259; PID:g1730260; PIDN:AAB38524.1
A;Experimental source: articular chondrocytes
C;Superfamily: aggrecan; C-type lectin homology; complement factor H repeat homology; I
C;Keywords: cartilage; chondroitin sulfate proteoglycan; extracellular matrix; glycopr

Query Match 14.2%; Score 158.5; DB 2; Length 2327;
Best Local Similarity 26.8%; Pred. No. 1.9e-05;
Matches 48; Conservative 29; Mismatches 59; Indels 43; Gaps 10;

QY 9 QPVCRRGG--TORPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSIIESEDEQKLIKFI 65
Db 2114 QKLCEEGWTKFQGHCHYR--HFPD---RATWVDAESQCRKQOSHLSIVTPEEQ----EFV 2164

QY 66 ENLLPSDGF-WIGLRRREEKQSNSTACQDLYAWTDGSIQFRNMYVDEP-----SCGSEV 120
Db 2165 NN---NAQDYQWIGL-----NDKTIEGDFRWSDGHSLOFENWRPNQDNFFATGEDC 2213

QY 121 CVVMYHQSAPAGIGGPFYMFQWNDRCNMKNPFICKYS-----DEKPAVPSREAEGETE 175
Db 2214 VVMIWHEKG-----EWNDVPCNYQLPFTCKKGTVACGEPVVEHARIFGQKKD 2261

RESULT 13
A39086
aggrecan precursor, cartilage long splice form [validated] - human
N;Alternate names: chondroitin sulfate proteoglycan 1; large aggregating proteoglycan;
N;Contains: aggrecan cartilage short splice form
C;Species: Homo sapiens (man)
C;Date: 10-Sep-1999 #sequence_revision 01-Dec-2000 #text_change 08-Dec-2000


```

F;126,239,333,387,611,667,1842/Binding site: carbohydrate (Asn) (covalent) #status pre

Query Match
Best Local Similarity 13.8%; Score 153.5; DB 2; Length 2124;
Matches 47; Conservative 27; Mismatches 62; Indels 43; Gaps 9;

QY 9 QPVCRGG---TQRPCYKVIYFHTSRRRLNFEAEKACRRDGGQLVSI ESEDEQKLI EKFI 65
Db 1911 QECC EEWTKFGHCHYR--HFPD---RETWVDAERRCREQQSHLSSIVTPEEQEFVNKNA 1965

QY 66 ENLLPSDGDF-WIGLRRREEKOSNSTACODLYAWTGDGSI SFRNWWYVDEP-----SCGSEV 120

```

Db	1966	Q	-----	DYQWIGL	-----	NDRTIEGDFRMSDGHSLQFEKWRPNQDNFFATGEDC	2010
Qv	121	CWVMYHOPAPAGIGGPPYMFOWNDDRCNMKNNEI	CKYS	----	DEKPAVPSPEAEGEETE	1175	

Db 2011 VVMITHERG-----EWNDVPCNYQLPFTCKKGTVACGEPVAVEHARTLGQKD 2058

Search completed: December 22, 2003, 16:14:40
Job time : 10.0151 secs

RESULT 15

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 22, 2003, 16:07:10 ; Search time 19.6848 Seconds
(without alignments)
1954.412 Million cell updates/sec

Title: US-09-887-855-2_COPY_22_227
Perfect score: 1115
Sequence: 1 ATGRLLSQGPVCRGGTQRPC.....EEDAKKTFKESREAAALNLAY 206

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 696363 seqs, 186758610 residues

Total number of hits satisfying chosen parameters: 696363

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :		Published Applications AA:*	
1:	/cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*	2:	/cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3:	/cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*	4:	/cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5:	/cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*	6:	/cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7:	/cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*	8:	/cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9:	/cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*	10:	/cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11:	/cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*	12:	/cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13:	/cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*	14:	/cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15:	/cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*	16:	/cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17:	/cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*	18:	/cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1115	100.0	206	9	US-09-887-855-5
2	1115	100.0	374	9	US-09-887-855-2
3	1115	100.0	374	12	US-10-094-749-2090
4	1115	100.0	374	15	US-10-149-819-15
5	1111	99.6	374	12	US-10-094-749-2142
6	1101	98.7	382	10	US-09-909-320-137
7	1101	98.7	382	10	US-09-909-088B-137
8	1101	98.7	382	10	US-09-905-291A-137
9	1101	98.7	382	10	US-09-902-853-137
10	1101	98.7	382	10	US-09-907-824-137
11	1101	98.7	382	10	US-09-907-841-137
12	1101	98.7	382	11	US-09-904-011-137
13	1101	98.7	382	11	US-09-906-742-137
14	1101	98.7	382	11	US-09-906-838-137
15	1101	98.7	382	11	US-09-907-613-137

Jack

16	1101	98.7	382	11	US-09-907-942-137	Sequence 137, App
17	1101	98.7	382	11	US-09-904-859-137	Sequence 137, App
18	1101	98.7	382	11	US-09-909-204-137	Sequence 137, App
19	1101	98.7	382	11	US-09-904-820-137	Sequence 137, App
20	1101	98.7	382	11	US-09-904-786-137	Sequence 137, App
21	1101	98.7	382	11	US-09-906-646-137	Sequence 137, App
22	1101	98.7	382	11	US-09-906-700-137	Sequence 137, App
23	1101	98.7	382	11	US-09-903-786-137	Sequence 137, App
24	1101	98.7	382	11	US-09-902-903-137	Sequence 137, App
25	1101	98.7	382	11	US-09-903-749A-137	Sequence 137, App
26	1101	98.7	382	11	US-09-904-119-137	Sequence 137, App
27	1101	98.7	382	11	US-09-904-956-137	Sequence 137, App
28	1101	98.7	382	11	US-09-902-736-137	Sequence 137, App
29	1101	98.7	382	11	US-09-907-794-137	Sequence 137, App
30	1101	98.7	382	11	US-09-903-943-137	Sequence 137, App
31	1101	98.7	382	11	US-09-904-462-137	Sequence 137, App
32	1101	98.7	382	11	US-09-907-925-137	Sequence 137, App
33	1101	98.7	382	11	US-09-902-692-137	Sequence 137, App
34	1101	98.7	382	11	US-09-903-520-137	Sequence 137, App
35	1101	98.7	382	11	US-09-905-056-137	Sequence 137, App
36	1101	98.7	382	11	US-09-909-064-137	Sequence 137, App
37	1101	98.7	382	11	US-09-904-553-137	Sequence 137, App
38	1101	98.7	382	11	US-09-905-381-137	Sequence 137, App
39	1101	98.7	382	11	US-09-905-088-137	Sequence 137, App
40	1101	98.7	382	11	US-09-907-575-137	Sequence 137, App
41	1101	98.7	382	11	US-09-905-075-137	Sequence 137, App
42	1101	98.7	382	11	US-09-902-759-137	Sequence 137, App
43	1101	98.7	382	11	US-09-902-634-137	Sequence 137, App
44	1101	98.7	382	11	US-09-902-713-137	Sequence 137, App
45	1101	98.7	382	11	US-09-907-979-137	Sequence 137, App

ALIGNMENTS

RESULT 1
US-09-887-855-5
; Sequence 5, Application US/09887855
; Patent No. US20020058310A1
; GENERAL INFORMATION:
; APPLICANT: Immunex Corporation
; APPLICANT: Anderson, Dirk M
; TITLE OF INVENTION: LECTIN SS3939 DNA AND POLYPEPTIDES
; FILE REFERENCE: 2883-US
; CURRENT APPLICATION NUMBER: US/09/887,855
; CURRENT FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 206
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-887-855-5

Query Match 100.0%; Score 1115; DB 9; Length 206;
Best Local Similarity 100.0%; Pred. No. 7.7e-105;
Matches 206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGRLLSQGPVCRGGTQRPCYKVIYFHDTSRRLNPFEEAKEACRRDGGQLVSI	60
Db	1	ATGRLLSQGPVCRGGTQRPCYKVIYFHDTSRRLNPFEEAKEACRRDGGQLVSI	60
QY	61	IEKFIENLLPSDGFDFWIGLRRREEKQSNSTACQDLYAWTDGSI	120
Db	61	IEKFIENLLPSDGFDFWIGLRRREEKQSNSTACQDLYAWTDGSI	120
QY	121	CVVMYHQSAPAGIGGPFYMFQWDDRCNMKNFICKYSDEKPAVPSREAGEETELTPV	180
Db	121	CVVMYHQSAPAGIGGPFYMFQWDDRCNMKNFICKYSDEKPAVPSREAGEETELTPV	180
QY	181	LPEETQEDAKKTFKESREAAALNLAY	206
Db	181	LPEETQEDAKKTFKESREAAALNLAY	206

Qy 121 CVVMYHQSAPAGIGGYPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGETELTPV 180
Db 142 CVVMYHQSAPAGIGGYPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGETELTPV 201
Qy 181 LPEETQEDAKTKFKESREAAALNLAY 206
Db 202 LPEETQEDAKTKFKESREAAALNLAY 227

RESULT 5

US-10-094-749-2142
; Sequence 2142, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2142
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-094-749-2142

Query Match 99.6%; Score 1111; DB 12; Length 374;
Best Local Similarity 99.5%; Pred. No. 4.4e-104;
Matches 205; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGRLLSGQVPCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSISEDEQKL 60
Db 22 ATGRLLSGQVPCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSISEDEQKL 81
Qy 61 IEKFIENLLPSDGDGFWIGLRRREEKQSNSTACQDLYAWTDGSGISQFRNYYVDEPSCGSEV 120
Db 82 IEKFIENLLPSDGDGFWIGLRRREEKQSNSTACQDLYAWTDGSGISQFRNYYVDEPSCGSEV 141
Qy 121 CVVMYHQSAPAGIGGYPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGETELTPV 180
Db 142 CVVMYHQSAPAGIGGYPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGETELTPV 201
Qy 181 LPEETQEDAKTKFKESREAAALNLAY 206
Db 202 LPEETQEDTKTKFKESREAAALNLAY 227

RESULT 6

US-09-909-320-137
; Sequence 137, Application US/09909320

; Patent No. US20020132240A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/909,320
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-909-320-137


```

; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/905,291A
; CURRENT FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-905-291A-137

Query Match      98.7%; Score 1101; DB 10; Length 382;
Best Local Similarity 96.3%; Pred. No. 4.6e-103;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 ATGRLLS-----GQVCRGGTQPCYKVIYFHDTSRRLNPFEEAKEACRRDGGQLVSI 52
Db 22 ATGRLLSASDLDLRGQGPVCRGGTQPCYKVIYFHDTSRRLNPFEEAKEACRRDGGQLVSI 81

QY 53 ESEDEQKLIKFIEIENLLPSDGFWIGLRRREKQSNSTACQDLYAWTDGSIQFRNWWYVD 112
Db 82 ESEDEQKLIKFIEIENLLPSDGFWIGLRRREKQSNSTACQDLYAWTDGSIQFRNWWYVD 141

QY 113 EPSCGSEVCVMYHQPSAPAGIGGPFYMFQWDDRCNKNFNICKYSDEKPAVPSREAEGE 172
Db 142 EPSCGSEVCVMYHQPSAPAGIGGPFYMFQWDDRCNKNFNICKYSDEKPAVPSREAEGE 201

QY 173 ETELTTPVLPEETQEEADAKTFKESREAAALNLAY 206
Db 202 ETELTTPVLPEETQEEADAKTFKESREAAALNLAY 235

RESULT 9
US-09-902-853-137
; Sequence 137, Application US/09902853
; Publication No. US20020192659A1
```

```

; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/902,853
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US/09/665,350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-902-853-137
```


Best Local Similarity 96.3%; Pred. No. 4.6e-103; Mismatches 0; Indels 8; Gaps 1;
Matches 206; Conservative 0;
QY 1 ATGRLLS-----GQVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
Db 22 ATGRLLSASDLDLRGGQPVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81
QY 53 ESEDEQKLIIEKFIENLLPSDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNWWYVD 112
Db 82 ESEDEQKLIIEKFIENLLPSDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNWWYVD 141
QY 113 EPSCGSEVCVMYHQPSAPAGIGGPFYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGE 172
Db 142 EPSCGSEVCVMYHQPSAPAGIGGPFYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGE 201
QY 173 ETELTPVLPEETQEEADAKKTFKESREAAALNLAY 206
Db 202 ETELTPVLPEETQEEADAKKTFKESREAAALNLAY 235

RESULT 10
US-09-907-824-137
; Sequence 137, Application US/09907824
; Publication No. US20020197671A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/907,824
; CURRENT FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: 09/665,350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15

; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-907-824-137
Query Match 98.7%; Score 1101; DB 10; Length 382;
Best Local Similarity 96.3%; Pred. No. 4.6e-103;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
QY 1 ATGRLLS-----GQVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
Db 22 ATGRLLSASDLDLRGGQPVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81
QY 53 ESEDEQKLIIEKFIENLLPSDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNWWYVD 112
Db 82 ESEDEQKLIIEKFIENLLPSDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNWWYVD 141
QY 113 EPSCGSEVCVMYHQPSAPAGIGGPFYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGE 172
Db 142 EPSCGSEVCVMYHQPSAPAGIGGPFYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGE 201
QY 173 ETELTPVLPEETQEEADAKKTFKESREAAALNLAY 206
Db 202 ETELTPVLPEETQEEADAKKTFKESREAAALNLAY 235

RESULT 11
US-09-907-841-137
; Sequence 137, Application US/09907841
; Publication No. US20020198366A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/907,841
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-907-841-137

Query Match 98.7%; Score 1101; DB 10; Length 382;
Best Local Similarity 96.3%; Pred. No. 4.6e-103;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 ATGRLLS-----GQVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
Db 22 ATGRLLSASDLRLGGQPVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

QY 53 ESEDEQKLIKFIENLLPSDGFQWIGLRRREKQSNSTACQDLYAWTDGSIQFRNYYVD 112
Db 82 ESEDEQKLIKFIENLLPSDGFQWIGLRRREKQSNSTACQDLYAWTDGSIQFRNYYVD 141

QY 113 EPSCGSEVCVMYHQPSAPAGIGGPFYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGE 172
Db 142 EPSCGSEVCVMYHQPSAPAGIGGPFYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGE 201

QY 173 ETELTPVLPEETQEDAKKTFKESREAAALNLAY 206
Db 202 ETELTPVLPEETQEDAKKTFKESREAAALNLAY 235

RESULT 12
US-09-904-011-137
; Sequence 137, Application US/09904011
; Publication No. US20030003530A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/904,011
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 09/665,350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-904-011-137

Query Match 98.7%; Score 1101; DB 11; Length 382;
Best Local Similarity 96.3%; Pred. No. 4.6e-103;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 ATGRLLS-----GQVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
Db 22 ATGRLLSASDLRLGGQPVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

QY 53 ESEDEQKLIKFIENLLPSDGFQWIGLRRREKQSNSTACQDLYAWTDGSIQFRNYYVD 112
Db 82 ESEDEQKLIKFIENLLPSDGFQWIGLRRREKQSNSTACQDLYAWTDGSIQFRNYYVD 141

QY 113 EPSCGSEVCVMYHQPSAPAGIGGPPYMFQWNDRCNMKNFICKYSDEKPAVPSREAE 172
Db 142 EPSCGSEVCVMYHQPSAPAGIGGPPYMFQWNDRCNMKNFICKYSDEKPAVPSREAE 201
QY 173 ETELTTPVLPEETQEDAKKTFKESREAAALNLAY 206
Db 202 ETELTTPVLPEETQEDAKKTFKESREAAALNLAY 235

RESULT 13
US-09-906-742-137
; Sequence 137, Application US/09906742
; Publication No. US20030023054A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/906,742
; CURRENT FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 09/665,350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02

; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-906-742-137

Query Match 98.7%; Score 1101; DB 11; Length 382;
Best Local Similarity 96.3%; Pred. No. 4.6e-103;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
QY 1 ATGRLLS-----GQVCRGGTQPCYKVIYFHDTSRRLNFEFEAKEACRRDGGQLVSI 52
Db 22 ATGRLLSASDLRLGGQPVCRGGTQPCYKVIYFHDTSRRLNFEFEAKEACRRDGGQLVSI 81
QY 53 ESEDEQKLIIEKFIENLLPSDGFWIGLRRREKQSNSTACQDLYAWTDGSIQFRNYYVD 112
Db 82 ESEDEQKLIIEKFIENLLPSDGFWIGLRRREKQSNSTACQDLYAWTDGSIQFRNYYVD 141
QY 113 EPSCGSEVCVMYHQPSAPAGIGGPPYMFQWNDRCNMKNFICKYSDEKPAVPSREAE 172
Db 142 EPSCGSEVCVMYHQPSAPAGIGGPPYMFQWNDRCNMKNFICKYSDEKPAVPSREAE 201
QY 173 ETELTTPVLPEETQEDAKKTFKESREAAALNLAY 206
Db 202 ETELTTPVLPEETQEDAKKTFKESREAAALNLAY 235

RESULT 14
US-09-906-838-137
; Sequence 137, Application US/09906838
; Publication No. US20030027143A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/906,838
; CURRENT FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 09/665,350
; PRIOR FILING DATE: 2000-09-18

; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-906-838-137

Query Match 98.7%; Score 1101; DB 11; Length 382;
Best Local Similarity 96.3%; Pred. No. 4.6e-103;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 ATGRLLS-----GQVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
Db 22 ATGRLLSASDLRLRGQPVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

QY 53 ESEDEQKLIKFIENLLPSDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNWYVD 112
Db 82 ESEDEQKLIKFIENLLPSDGFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNWYVD 141

QY 113 EPSCGSEVCVMYHQPAPAGIGGPPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGE 172
Db 142 EPSCGSEVCVMYHQPAPAGIGGPPYMFQWDDRCNMKNFICKYSDEKPAVPSREAEGE 201

QY 173 ETELTPVLPEETOEDAKKTFKESREAAALNLAY 206
Db 202 ETELTPVLPEETOEDAKKTFKESREAAALNLAY 235

RESULT 15
US-09-907-613-137
; Sequence 137, Application US/09907613
; Publication No. US20030027145A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone

; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/907,613
; PRIOR FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-907-613-137

Query Match 98.7%; Score 1101; DB 11; Length 382;
Best Local Similarity 96.3%; Pred. No. 4.6e-103;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 ATGRLLS-----GQVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 52
Db 22 ATGRLLSASDLRLRGQPVCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDGGQLVSI 81

[illegible]

Search completed: December 22, 2003, 16:16:47
Job time : 20.6848 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 22, 2003, 16:05:09 ; Search time 11.9128 Seconds
(without alignments)
1476.391 Million cell updates/sec

Title: US-09-887-855-2

Perfect score: 2000

Sequence: 1 MRPQTALQAVLLVLLVGLR.....PDQMGRSKESGWENEIYGY 374

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	657	32.9	273	1	CHOD_MOUSE
2	645	32.2	273	1	CHOD_HUMAN
3	183	9.2	1456	1	MANR_HUMAN
4	177	8.8	1268	1	PGCN_MOUSE
5	176	8.8	1321	1	PGCN_HUMAN
6	174.5	8.7	3381	1	PGCV_BOVIN
7	174	8.7	1257	1	PGCN_RAT
8	174	8.7	2738	1	PGCV_RAT
9	174	8.7	3358	1	PGCV_MOUSE
10	174	8.7	3396	1	PGCV_HUMAN
11	172.5	8.6	643	1	CD93_RAT
12	171	8.6	3562	1	PGCV_CHICK
13	158.5	7.9	2364	1	PGCA_BOVIN
14	158.5	7.9	2415	1	PGCA_HUMAN
15	157.5	7.9	652	1	CD93_HUMAN
16	155.5	7.8	2333	1	PGCA_CANFA
17	154.5	7.7	612	1	LEM2_MOUSE
18	154.5	7.7	644	1	CD93_MOUSE
19	153.5	7.7	158	1	LECG_TRIST
20	153.5	7.7	2124	1	PGCA_RAT
21	153	7.6	162	1	LEC3_MEGRO
22	152	7.6	912	1	PGCB_BOVIN
23	151	7.5	197	1	CLF1_HUMAN
24	149	7.4	321	1	FCE2_HUMAN
25	148.5	7.4	152	1	IXA_TRIFL
26	148.5	7.4	2132	1	PGCA_MOUSE
27	147	7.3	883	1	PGCB_MOUSE
28	146	7.3	2109	1	PGCA_CHICK
29	145	7.2	883	1	PGCB_RAT
30	144.5	7.2	173	1	LEC2_MEGRO
31	144.5	7.2	372	1	LEM1_RAT
32	143.5	7.2	372	1	LEM1_MOUSE
33	143	7.1	549	1	LEM2_RAT
					Q9cxm0 mus musculu
					Q9h9p2 homo sapien
					P22897 homo sapien
					P55066 mus musculu
					O14594 homo sapien
					P81282 bos taurus
					P55067 rattus norv
					Q9erb4 rattus norv
					P62059 mus musculu
					P13611 homo sapien
					Q9et61 rattus norv
					Q90953 gallus gall
					P13608 bos taurus
					P16112 homo sapien
					Q9npy3 homo sapien
					Q28343 canis famil
					Q00690 mus musculu
					Q89103 mus musculu
					Q9y9p1 trimeresuru
					P07897 rattus norv
					P07439 megabalanus
					Q28062 bos taurus
					P06734 homo sapien
					P23806 trimeresuru
					Q61282 mus musculu
					Q61361 mus musculu
					P07898 gallus gall
					P55068 rattus norv
					P17346 megabalanus
					P30836 rattus norv
					P18337 mus musculu
					P98105 rattus norv

34	141.5	7.1	165	1	LIT1_MOUSE	P43137 mus musculu
35	141.5	7.1	331	1	FCE2_MOUSE	P20693 mus musculu
36	140	7.0	370	1	LEM1_BOVIN	P98131 bos taurus
37	139	7.0	248	1	PSPA_HUMAN	P07714 homo sapien
38	139	7.0	283	1	LECA_SARPE	P05047 sarcophaga
39	138.5	6.9	172	1	LECA_PLEWA	Q02988 pleurodeles
40	138.5	6.9	202	1	TETN_MOUSE	P43025 mus musculu
41	137.5	6.9	372	1	LEM1_MACMU	Q95198 macaca mula
42	137.5	6.9	372	1	LEM1_PAPHA	Q28768 papio hamad
43	136	6.8	175	1	LITH_BOVIN	P23132 bos taurus
44	134.5	6.7	132	1	ACAL_ANSAN	P83300 anser anser
45	134.5	6.7	372	1	LEM1_PONPY	Q95235 pongo pygma

ALIGNMENTS

RESULT 1
CHOD_MOUSE
ID CHOD_MOUSE STANDARD; PRT; 273 AA.
AC Q9CXMO; Q8VI31;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Chondrolectin precursor (Transmembrane protein MT75).
GN CHODL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RA Weng L., Smits P., Hubner R., Wouters J., Merregaert J.;
RT "Mt75, a low expressed c-type lectin gene involving in
RT chondrogenesis";
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryonic head;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -|- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -|- SIMILARITY: Contains 1 C-type lectin family domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AF311699; AAL50354.1; -.

DR EMBL; AK014255; BAB29226.1; -.
DR HSSP; P22897; 1EGG.
DR MGD; MGI:2179069; Chodl.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00059; lectin_c; 1.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C_TYPE_LECTIN_1; FALSE_NEG.
DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW Lectin; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 273 CHONDROLECTIN.
FT DOMAIN 22 216 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 217 237 POTENTIAL.
FT DOMAIN 238 273 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 35 179 C-TYPE LECTIN.
FT CARBOHYD 86 86 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 24 24 V -> W (IN REF. 2).
FT CONFLICT 179 179 T -> K (IN REF. 2).
SQ SEQUENCE 273 AA; 30303 MW; E052D933F244F4C7 CRC64;

Query Match 32.9%; Score 657; DB 1; Length 273;
Best Local Similarity 49.1%; Pred. No. 3.4e-48;
Matches 139; Conservative 38; Mismatches 80; Indels 26; Gaps 9;

QY 10 VLLAVLLVGLRAATGRLLSGQPVCRGGTORPCYKVIYFHDTSRRLLNFEEAKEACRRDGGQ 69
Db 8 LLGAALLCAQGAFAFRRVVGQKCFADVKHPCYKMYAFHELSSRVSFQEARLACESEGGV 67

QY 70 LVSISEDEQKLIKFIENLLP-----SDGDFWIGLRRREEKQSNSTACQDLYAWTDGSI 124
Db 68 LLSLENEAEQKLIESMLQNLTKPGTGISDGDWIGLRSRGDGT-SGACPDLYQWSDGSS 126

QY 125 SQFRNYYVDEPSCGSEVCVVMYHQPSAPAGIGGYPYFQWDDRCNMKNFICKYSDE-KP 183
Db 127 SQFRNYYTDEPSCGSEKCVVMYHQPTANPGLGGPYLYQWDDRCNMKNHYICTYEPEIHP 186

QY 184 AVPSREAGEETEELTPVLPEETOEDAKKTFKESREAAAL--NLAYILIPSLPLLLLVV 241
Db 187 TEPA-----EKPYLTNQ--PEETHENVV-----VTEAGIIPNLIYVIPTIPLLLILV 233

QY 242 TTVCVWVICRKRK-REQDPDSTKKQHTIWPSPHOQNSPDLEV 283
Db 234 ALGTCCFQMLHKSGRSKTSPN---QSTLWISKSTRKESGMEV 273

RESULT 2
CHOD_HUMAN STANDARD; PRT; 273 AA.
AC Q9H9P2; Q9HCY3;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Chondrolectin precursor (Transmembrane protein MT75) (PRED12 protein).
DE CHODL OR C21ORF68.
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RX MEDLINE=22074930; PubMed=12079284;
RA Weng L., Smits P., Wauters J., Merregaert J.;
RT "Molecular cloning and characterization of human chondrolectin, a novel type I transmembrane protein homologous to C-type lectins.";
RL Genomics 80:62-70(2002).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20289799; PubMed=10830953;
RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T., Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Groner Y., Soeda E., Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A., Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,

RA Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W., Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S., Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E., Minoshima S., Shimizu N., Nordsiek G., Hornischer K., Brandt P., Scharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloecker H., Ramser J., Beck A., Klages S., Hennig S., Riesselmann L., Dagand E., Wehrmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F., Lehrach H., Reinhardt R., Yaspo M.-L.;
RT "The DNA sequence of human chromosome 21.";
RL Nature 405:311-319(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP SEQUENCE OF 27-273 FROM N.A.
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y., Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H., Wagatsuma M., Hagoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M., Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S., Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y., Ninomiya K., Iwayanagi T.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP TISSUE SPECIFICITY.
RX MEDLINE=21564202; PubMed=11707072;
RA Raymond A., Friedli M., Neergard Henriksen C., Chapot F., Deutsch S., Ucla C., Rossier C., Lyle R., Guipponi M., Antonarakis S.E.;
RT "From PREDs and open reading frames to cDNA isolation: revisiting the human chromosome 21 transcription map.";
RL Genomics 78:46-54(2001).
CC -I- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -I- TISSUE SPECIFICITY: Found in spleen, testis, prostate and fetal liver. Expression limited to vascular muscle of testis, smooth muscle of prostate stroma, heart muscle, skeletal muscle, crypts of small intestine, and red pulp of spleen.
CC -I- PTM: N-glycosylated.
CC -I- SIMILARITY: Contains 1 C-type lectin family domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF257472; AAL05981.1; -
CC EMBL; AL163217; CAB90388.1; -
CC EMBL; BC009418; AAH09418.1; -
CC EMBL; AK022689; BAB14181.1; ALT_INIT.

DR	InterPro; IPR003006; Ig_MHC.	
DR	InterPro; IPR001304; Lectin_C.	
DR	InterPro; IPR000538; Link.	
DR	InterPro; IPR000436; Sushi_SCR_CCP.	
DR	Pfam; PF00008; EGF; 2.	
DR	Pfam; PF00047; Ig; 1.	
DR	Pfam; PF00059; lectin_c; 1.	
DR	Pfam; PF00084; sushi; 1.	
DR	Pfam; PF00193; Xlink; 2.	
DR	PRINTS; PR01265; LINKMODULE.	
DR	PRINTS; PR00356; ANTIFREEZEII.	
DR	ProDom; PD000918; Link; 2.	
DR	SMART; SM00032; CCP; 1.	
DR	SMART; SM00034; CLECT; 1.	
DR	SMART; SM00179; EGF_CA; 1.	
DR	SMART; SM00409; IG; 1.	
DR	SMART; SM00445; LINK; 2.	
DR	PROSITE; PS00010; ASX_HYDROXYL; 1.	
DR	PROSITE; PS00615; C_TYPE_LECTIN_1; 1.	
DR	PROSITE; PS50041; C_TYPE_LECTIN_2; 1.	
DR	PROSITE; PS00022; EGF_1; 3.	
DR	PROSITE; PS01186; EGF_2; 1.	
DR	PROSITE; PS01187; EGF_CA; 1.	
DR	PROSITE; PS50835; IG_LIKE; 1.	
DR	PROSITE; PS01241; LINK; 2.	
KW	Glycoprotein; Hyaluronic acid; Proteoglycan; Immunoglobulin domain;	
KW	EGF-like domain; Calcium; Repeat; Sushi; Signal.	
FT	SIGNAL 1 22	
FT	CHAIN 23 1268	
FT	DOMAIN 37 157	
FT	DOMAIN 158 253	
FT	DOMAIN 259 355	
FT	DOMAIN 960 996	
FT	DOMAIN 998 1034	
FT	DOMAIN 1036 1165	
FT	DOMAIN 1166 1224	
FT	DISULFID 58 139	
FT	DISULFID 181 252	
FT	DISULFID 205 226	
FT	DISULFID 279 354	
FT	DISULFID 303 324	
FT	DISULFID 964 975	
FT	DISULFID 969 984	
FT	DISULFID 986 995	
FT	DISULFID 1040 1051	
FT	DISULFID 1068 1160	
FT	DISULFID 1136 1152	
FT	DISULFID 1167 1210	
FT	DISULFID 1196 1223	
FT	CARBOHYD 121 121	(POTENTIAL).
FT	CARBOHYD 339 339	(POTENTIAL).
FT	CARBOHYD 742 742	(POTENTIAL).
FT	CARBOHYD 978 978	(POTENTIAL).
FT	CARBOHYD 1175 1175	(POTENTIAL).
SQ	SEQUENCE 1268 AA; 137200 MW; 3014E8E202A2FAEC CRC64;	
Query Match 8.8%; Score 177; DB 1; Length 1268;		
Best Local Similarity 31.5%; Pred. No. 1e-06;		
Matches 45; Conservative 17; Mismatches 49; Indels 32; Gaps 6;		
QY	38 QRPCYKVIYFHTDTRRLNPFEEAKEACRRDGGQLVSIIESEDEQKLEKFIENLLPSDGDWF 97	
Db	1048 QGHYR--YF---AHRRAWEADAERDCRRRAGHLTSVHSPEEHKFINSF-----GHENSW 1096	
QY	98 IGLRRREEKQSNSTACQDLYAWTDGSIQFRNYYVDEPS---CGSEVCVMYHQPSPAPAG 154	
Db	1097 IGLNDRTVRD-----FQWTDNTGLQYENWEKQPDNFFAGGEDCVMVMAHESG--- 1145	
QY	155 IGGPYMFQWNDRCNMKNPFICK 177	
Db	1146 -----RWNDVPCNPNLPYVCK 1161	

RESULT 5	
PGCN HUMAN	
ID -PGCN HUMAN STANDARD; PRT; 1321 AA.	
AC O14594; Q9UPK6;	
DT 28-FEB-2003 (Rel. 41, Created)	
DT 28-FEB-2003 (Rel. 41, Last sequence update)	
DT 15-SEP-2003 (Rel. 42, Last annotation update)	
DE Neurocan core protein precursor (Chondroitin sulfate proteoglycan 3).	
GN CSPG3 OR NCAN OR NEUR.	
OS Homo sapiens (Human).	
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;	
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
OX NCBI_TaxID=9606;	
RN [1]	
RP SEQUENCE FROM N.A.	
RX MEDLINE=99013874; PubMed=9795216;	
RA Prange C.K., Pennacchio L.A., Lieuallen K., Fan W., Lennon G.G.;	
RT "Characterization of the human neurocan gene, CSPG3.";	
RL Gene 221:199-205(1998).	
RN [2]	
RP SEQUENCE OF 1-990 AND 1007-1321 FROM N.A.	
RA Lamerdin J.E., McCready P.M., Skowronski E., Adamson A.W.,	
RA Burkhardt-Schultz K., Gordon L., Kyle A., Ramirez M., Stilwagen S.,	
RA Phan H., Velasco N., Do L., Regala W., Terry A., Garnes J.,	
RA Danganan L., Poundstone P., Christensen M., Georgescu A., Avila J.,	
RA Liu S., Attix C., Andreise T., Frankheim M., Amico-Keller G.,	
RA Coefield J., Duarte S., Lucas S., Bruce R., Thomas P., Quan G.,	
RA Kronmiller B., Arellano A., Montgomery M., Ow D., Nolan M., Trong S.,	
RA Kobayashi A., Olsen A.S., Carrano A.V.;	
RT "Sequence analysis of an ~1 Mb region containing the MEF2B gene in	
RT 19p12.";	
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.	
CC -!- FUNCTION: May modulate neuronal adhesion and neurite growth during	
CC development by binding to neural cell adhesion molecules (NG-CAM	
CC and N-CAM). Chondroitin sulfate proteoglycan; binds to hyaluronic	
CC acid.	
CC -!- TISSUE SPECIFICITY: Brain.	
CC -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.	
CC -!- SIMILARITY: Contains 2 EGF-like domains.	
CC -!- SIMILARITY: Contains 2 link domains.	
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.	
CC -!- SIMILARITY: Contains 1 Sushi (SCR) domain.	
CC -!- SIMILARITY: BELONGS TO THE AGGREGAN/VERSICAN PROTEOGLYCAN FAMILY.	

This SWISS-PROT entry is copyright. It is produced through a collaboration	
between the Swiss Institute of Bioinformatics and the EMBL outstation -	
the European Bioinformatics Institute. There are no restrictions on its	
use by non-profit institutions as long as its content is in no way	
modified and this statement is not removed. Usage by and for commercial	
entities requires a license agreement (See http://www.isb-sib.ch/announce/	
or send an email to license@isb-sib.ch).	

EMBL; AF026547; AAC80576.1; -	
EMBL; AC003110; AAB86655.1; -	
EMBL; AC005254; AAC25581.1; -	
HSSP; P00740; 1EDM.	
Genew; HGNC:2465; CSPG3.	
MIM; 600826; -	
InterPro; IPR000152; Asx_hydroxyl.	
InterPro; IPR000742; EGF_2.	
InterPro; IPR001881; EGF_Ca.	
InterPro; IPR006209; EGF_like.	
InterPro; IPR007110; Ig-Like.	
InterPro; IPR003599; Ig.	
InterPro; IPR003006; Ig_MHC.	
InterPro; IPR001304; Lectin_C.	
InterPro; IPR000538; Link.	
InterPro; IPR000436; Sushi_SCR_CCP.	
Pfam; PF00008; EGF; 2.	
Pfam; PF00047; Ig; 1.	
Pfam; PF00059; lectin_c; 1.	
Pfam; PF00084; sushi; 1.	
Pfam; PF00193; Xlink; 2.	

```
DR PRINTS; PR01265; LINKMODULE.
DR ProDom; PD000918; Link; 2.
DR SMART; SM00032; CCP; 1.
DR SMART; SM00034; CLECT; 1.
DR SMART; SM00179; EGF_CA; 1.
DR SMART; SM00409; IG_1.
DR SMART; SM00445; LINK; 2.
DR PROSITE; PS00010; ASX_HYDROXYL; 1.
DR PROSITE; PS00615; C_TYPE_LLECTIN_1; 1.
DR PROSITE; PS00041; C_TYPE_LLECTIN_2; 1.
DR PROSITE; PS00022; EGF_1; 3.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01187; EGF_CA; 1.
DR PROSITE; PS01187; EGF_CA; 1.
DR PROSITE; PS01241; LINK; 2.
DR Glycoprotein; Hyaluronic acid; Proteoglycan; Immunoglobulin domain;
KW EGF-like domain; Calcium; Repeat; Lentin; Sushi; Signal.
FT SIGNAL 1 22
FT CHAIN 23 1321
FT DOMAIN 38 153
FT DOMAIN 159 254
FT DOMAIN 260 356
FT DOMAIN 1008 1044
FT DOMAIN 1046 1082
FT DOMAIN 1084 1213
FT DOMAIN 1214 1272
FT DISULFID 59 140
FT DISULFID 182 253
FT DISULFID 206 227
FT DISULFID 280 355
FT DISULFID 304 325
FT DISULFID 1012 1023
FT DISULFID 1017 1032
FT DISULFID 1034 1043
FT DISULFID 1088 1099
FT DISULFID 1116 1208
FT DISULFID 1184 1200
FT DISULFID 1215 1258
FT DISULFID 1244 1271
FT CARBOHYD 122 122
FT CARBOHYD 340 340
FT CARBOHYD 1026 1026
FT CARBOHYD 1223 1223
FT CONFLICT 1234 1234
FT CONFLICT 1254 1254
FT CONFLICT 1282 1282
SQ SEQUENCE 1321 AA; 142972 MW; 2EF47F823DB980B8 CRC64;

Query Match 8.8%; Score 176; DB'1; Length 1321;
Best Local Similarity 31.5%; Pred. No. 1.3e-06;
Matches 45; Conservative 18; Mismatches 48; Indels 32; Gaps 6;

QY 38 QRPCYKVIYFHDTSRRLNFEAEKACRRDGGQGLVSI ESEDEQKLIKFIENLLPSDGDWF 97
Db 1096 QGHCYR--YF---AHRRAWEDAEDKDCRRSGHLTSVHSPEEHSFINSF-----GHENTW 1144

QY 98 IGLRRREKQSNSTACQDLYAWTDGSIQFRNWWYVDEPS---CGSEVCVVMYHQPSAPAG 154
Db 1145 IGLNDRIVERD-----FQWTDNTGLQFENWRENQPDNFFAGGEDCVVMVAHESG--- 1193

QY 155 IGGPYMFQWDDRCNMKNFICK 177
Db 1194 -----RWNDVPCYNLPYVCK 1209

RESULT 6
PGCV_BOVIN
ID PGCV_BOVIN STANDARD; PRT; 3381 AA.
AC P81282; O77609; O77610; O77611; O77612;
DT 15-DEC-1998 (Rel. 37, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Versican core protein precursor (Large fibroblast proteoglycan)
```

```
DE (Chondroitin sulfate proteoglycan core protein 2) (PG-M) (Glial
DE hyaluronate-binding protein) (GHAP).
GN CPBG2.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS VO; V1; V2 AND V3).
RC TISSUE=Forebrain;
RX MEDLINE=98288320; PubMed=9624174;
RA Schmalfeldt M., Dours-Zimmermann M.T., Winterhalter K.H.,
RA Zimmermann D.R.;
RT "Versican V2 is a major extracellular matrix component of the mature
RT bovine brain.";
RL J. Biol. Chem. 273:15758-15764(1998).
RN [2]
RP SEQUENCE OF 21-53; 78-96; 226-250; 262-277; 295-306; 314-324; 329-331
RP AND 342-348.
RC TISSUE=Spinal cord;
RX MEDLINE=92062692; PubMed=1720020;
RA Perides G., Biviano F., Bignami A.;
RT "Interaction of a brain extracellular matrix protein with hyaluronic
RT acid.";
RL Biochim. Biophys. Acta 1075:248-258(1991).
CC -!- FUNCTION: May play a role in intercellular signaling and in
CC connecting cells with the extracellular matrix. May take part in
CC the regulation of cell motility, growth and differentiation. Binds
CC hyaluronic acid.
CC -!- SUBUNIT: Interacts with FBLN1 (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=4;
CC Comment=Additional isoforms seem to exist;
CC Name=V0;
CC IsoId=P81282-1; Sequence=Displayed;
CC Name=V1;
CC IsoId=P81282-2; Sequence=VSP_003078, VSP_003079;
CC Name=V2;
CC IsoId=P81282-3; Sequence=VSP_003080;
CC Name=V3;
CC IsoId=P81282-4; Sequence=VSP_003078, VSP_003081;
CC -!- TISSUE SPECIFICITY: Cerebral white matter. V0 and V1 are expressed
CC in the central nervous system, and in a number of mesenchymal and
CC epithelial tissues; the major isoform V2 is restricted to the
CC central nervous system.
CC -!- DEVELOPMENTAL STAGE: Disappears after the cartilage development
CC (By similarity).
CC -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -!- SIMILARITY: Contains 2 link domains.
CC -!- SIMILARITY: Contains 2 EGF-like domains.
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC -!- SIMILARITY: Contains 1 Sushi (SCR) domain.
CC -!- SIMILARITY: BELONGS TO THE AGGRECAN/VERSICAN PROTEOGLYCAN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF060456; AAC24358.1; -.
CC EMBL; AF060457; AAC24359.1; -.
CC EMBL; AF060458; AAC24360.1; -.
CC EMBL; AF060459; AAC24361.1; -.
CC PIR; T14274; T14274.
CC PIR; T42389; T42389.
CC HSSP; P01132; LEPG.
CC InterPro; IPR000152; Asx_hydroxyl.
CC InterPro; IPR000742; EGF_2.
```

DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR007110; Ig-Like.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR000538; Link.
DR InterPro; IPR000436; Sushi_SCR_CCP.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF00059; lectin_c; 1.
DR Pfam; PF00084; sushi; 1.
DR Pfam; PF00193; xlink; 2.
DR PRINTS; PR01265; LINKMODULE.
DR PRINTS; PR00356; ANTIFREEZEII.
DR ProDom; PD000918; Link; 2.
DR SMART; SM00032; CCP; 1.
DR SMART; SM00034; CLECT; 1.
DR SMART; SM00179; EGF_CA; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00445; LINK; 2.
DR PROSITE; PS00010; ASX_HYDROXYL; 1.
DR PROSITE; PS00615; C_TYPE_LLECTIN_1; 1.
DR PROSITE; PS00441; C_TYPE_LLECTIN_2; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01187; EGF_CA; 1.
DR PROSITE; PS00835; IG_LIKE; 1.
DR PROSITE; PS01241; LINK; 2.
KW Glycoprotein; Proteoglycan; Lectin; Extracellular matrix; Sushi;
KW Signal; Repeat; EGF-like domain; Calcium; Immunoglobulin domain;
KW Hyaluronic acid; Alternative splicing.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 3381 VERSICAN CORE PROTEIN.
FT DOMAIN 21 147 IG-LIKE V-TYPE.
FT DOMAIN 168 245 LINK 1.
FT DOMAIN 266 347 LINK 2.
FT DOMAIN 349 1336 GAG-ALPHA
(GLUCOSAMINOGLYCAN ATTACHMENT DOMAIN).
FT DOMAIN 1337 3074 GAG-BETA.
FT DOMAIN 3074 3110 EGF-LIKE 1.
FT DOMAIN 3112 3148 EGF-LIKE 2, CALCIUM-BINDING.
FT DOMAIN 3161 3275 C-TYPE LECTIN.
FT DOMAIN 3280 3338 SUSHI.
FT DISULFID 44 131 BY SIMILARITY.
FT DISULFID 173 244 BY SIMILARITY.
FT DISULFID 197 218 BY SIMILARITY.
FT DISULFID 271 346 BY SIMILARITY.
FT DISULFID 295 316 BY SIMILARITY.
FT DISULFID 3078 3089 BY SIMILARITY.
FT DISULFID 3083 3098 BY SIMILARITY.
FT DISULFID 3100 3109 BY SIMILARITY.
FT DISULFID 3116 3127 BY SIMILARITY.
FT DISULFID 3121 3136 BY SIMILARITY.
FT DISULFID 3138 3147 BY SIMILARITY.
FT DISULFID 3154 3165 BY SIMILARITY.
FT DISULFID 3182 3274 BY SIMILARITY.
FT DISULFID 3250 3266 BY SIMILARITY.
FT DISULFID 3281 3324 BY SIMILARITY.
FT DISULFID 3310 3337 BY SIMILARITY.
FT CARBOHYD 57 57 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 331 331 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 817 817 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 965 965 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1017 1017 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1333 1333 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1393 1393 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1437 1437 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1463 1463 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1653 1653 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1974 1974 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2045 2045 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 2074 2074 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2103 2103 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2263 2263 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2290 2290 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2356 2356 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2623 2623 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2641 2641 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2919 2919 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 3052 3052 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 3354 3354 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 3364 3364 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 349 349 P -> R (in isoform V1 and isoform V3).
FT VARSPLIC 350 1336 /FTid=VSP 003078.
FT VARSPLIC 1337 3074 Missing (in isoform V1).
FT VARSPLIC 350 3074 Missing (in isoform V2).
FT VARSPLIC 350 3074 /FTid=VSP 003080.
FT VARSPLIC 350 3074 Missing (in isoform V3).
FT CONFLICT 25 25 /FTid=VSP 003081.
FT CONFLICT 51 51 MISSING (IN REF. 2).
FT CONFLICT 89 89 MISSING (IN REF. 2).
FT CONFLICT 96 96 N -> D (IN REF. 2).
FT CONFLICT 346 346 Q -> D (IN REF. 2).
SQ SEQUENCE 3381 AA; 369984 MW; F09716FA778D459 CRC64;
Query Match 8.7%; Score 174.5; DB 1; Length 3381;
Best Local Similarity 25.5%; Pred. No. 6.2e-06;
Matches 59; Conservative 26; Mismatches 71; Indels 75; Gaps 11;
QY 38 QRPCYKVIYFHDTSRRRLNFEAEKACRRDGGQLVSIIESEDEQKLIKFIENLLPSDGF- 96
Db 3162 QGQCYK--YF---AHRRTWDAARECRLOGAHLTSILSHEEQMFVNRV-----GHDYQ 3209
QY 97 WIGLRRREEKQSNSTACQDLYAWTDGSIQFRNRYVDEP-----SCGSEVCVMYHQSAP 152
Db 3210 WIGL-----NDKMFHDFRWTGSLQYENWRPNQDPSFFSTGEDCVVIWHENG--- 3259
QY 153 AGIGGPFYMFQWDDRCNMKNFICKYS-----DEKPAVPSREAEGE----- 193
Db 3260 -----QWNDVPCNYHLTYTCKGTGTVACGQPPWVENAKTFGKMKPRYEINSLIRYHC 3310
QY 194 -----ETELTT-----PVL-----PEETQEEDAKKTFKESREAAALN 224
Db 3311 KDGFIQRHLPTIRCLNGRWAMPKITCLNPSAYQRTYSKKYFKNSSSAKDN 3361
RESULT 7
ID_PGCN_RAT STANDARD; PRT; 1257 AA.
AC P55067;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Neurocan core protein precursor (Chondroitin sulfate proteoglycan 3)
DE (245 kDa early postnatal core glycoprotein) [Contains: 150 kDa adult
core glycoprotein].
GN CSPG3 OR NCAN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=Sprague-Dawley; TISSUE=Brain;
RX MEDLINE=92406907; PubMed=1326557;
RA Rauch U., Karthikeyan L., Maurel P., Margolis R.U., Margolis R.K.;
RT "Cloning and primary structure of neurocan, a developmentally
regulated, aggregating chondroitin sulfate proteoglycan of brain.";
RL J. Biol. Chem. 267:19536-19547(1992).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=94230574; PubMed=7513709;

RA Friedlander D.R., Milev P., Karthikeyan L., Margolis R.K.,
RA Margolis R.U., Grunet M.;
RT "The neuronal chondroitin sulfate proteoglycan neurocan binds to the
RT neural cell adhesion molecules Ng-CAM/L1/NILE and N-CAM, and inhibits
RT neuronal adhesion and neurite outgrowth.";
RL J. Cell Biol. 125:669-680(1994).
CC -!- FUNCTION: May modulate neuronal adhesion and neurite growth during
CC development by binding to neural cell adhesion molecules (NG-CAM
CC and N-CAM). Chondroitin sulfate proteoglycan; binds to hyaluronic
CC acid.
CC
CC -!- TISSUE SPECIFICITY: EARLY POSTNATAL AND ADULT BRAIN; NOT EXPRESSED
CC IN KIDNEY, LUNG, LIVER AND MUSCLE.
CC
CC -!- PTM: CONTAINS MOSTLY CHONDROITIN SULFATE, BUT ALSO N-LINKED AND
CC O-LINKED OLIGOSACCHARIDES (BY SIMILARITY).
CC
CC -!- PTM: TWO ISOFORMS WERE FOUND THAT PROBABLY ARISE BY PROTEOLYTIC
CC DEGRADATION. THE LARGE ISOFORM IS PREDOMINANT IN EARLY POSTNATAL
CC BRAIN, THE SMALL ISOFORM IS FOUND IN ADULT BRAIN.
CC
CC -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC
CC -!- SIMILARITY: Contains 2 EGF-like domains.
CC
CC -!- SIMILARITY: Contains 2 link domains.
CC
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC
CC -!- SIMILARITY: Contains 1 Sushi (SCR) domain.
CC
CC -!- SIMILARITY: BELONGS TO THE AGGRECAN/VERSICAN PROTEOGLYCAN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; M97161; AAC37679.1; -.
CC PIR; S28764; S28764.
CC HSSP; P00740; 1EDM.
CC
CC InterPro; IPR002353; AntifreezeII.
CC InterPro; IPR000152; Asx_hydroxyl.
CC InterPro; IPR000742; EGF_2.
CC InterPro; IPR001881; EGF_Ca.
CC InterPro; IPR006209; EGF_like.
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR003599; Ig.
CC InterPro; IPR003006; Ig_MHC.
CC InterPro; IPR001304; Lectin_C.
CC InterPro; IPR000538; Link.
CC InterPro; IPR000436; Sushi_SCR_CCP.
CC Pfam; PF00008; EGF; 2.
CC Pfam; PF00047; ig; 1.
CC Pfam; PF00059; lectin_c; 1.
CC Pfam; PF00084; sushi; 1.
CC Pfam; PF00193; Xlink; 2.
CC PRINTS; PR01265; LINKMODULE.
CC PRINTS; PR00356; ANTIFREEZEII.
CC ProDom; PD000918; Link; 2.
CC SMART; SM00032; CCP; 1.
CC SMART; SM00034; CLECT; 1.
CC SMART; SM00179; EGF_CA; 1.
CC SMART; SM00409; IG; 1.
CC SMART; SM00445; LINK; 2.
CC PROSITE; PS00010; ASX_HYDROXYL; 1.
CC PROSITE; PS00615; C_TYPE_LLECTIN_1; 1.
CC PROSITE; PS50041; C_TYPE_LLECTIN_2; 1.
CC PROSITE; PS00022; EGF_1; 3.
CC PROSITE; PS01186; EGF_2; 1.
CC PROSITE; PS01187; EGF_CA; 1.
CC PROSITE; PS50835; IG_LIKE; 1.
CC PROSITE; PS01241; LINK; 2.
KW Glycoprotein; Hyaluronic acid; Proteoglycan; Immunoglobulin domain;
KW EGF-like domain; Calcium; Repeat; Lectin; Sushi; Signal.
FT SIGNAL 1 22
FT CHAIN 23 1257 NEUROCAN CORE PROTEIN.
FT CHAIN 639 1257 150 kDa ADULT CORE GLYCOPROTEIN.
FT DOMAIN 37 157 IG-LIKE V-TYPE.

FT DOMAIN 158 253 LINK 1.
FT DOMAIN 259 355 LINK 2.
FT DOMAIN 949 985 EGF-LIKE 1.
FT DOMAIN 987 1023 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 1025 1154 C-TYPE LECTIN.
FT DOMAIN 1155 1213 SUSHI.
FT DISULFID 58 139 BY SIMILARITY.
FT DISULFID 181 252 BY SIMILARITY.
FT DISULFID 205 226 BY SIMILARITY.
FT DISULFID 279 354 BY SIMILARITY.
FT DISULFID 303 324 BY SIMILARITY.
FT DISULFID 953 964 BY SIMILARITY.
FT DISULFID 958 973 BY SIMILARITY.
FT DISULFID 975 984 BY SIMILARITY.
FT DISULFID 1029 1040 BY SIMILARITY.
FT DISULFID 1057 1149 BY SIMILARITY.
FT DISULFID 1125 1141 BY SIMILARITY.
FT DISULFID 1156 1199 BY SIMILARITY.
FT DISULFID 1185 1212 BY SIMILARITY.
FT CARBOHYD 121 121 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 339 339 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 737 737 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 944 944 O-LINKED (XYL. .) (CHONDROITIN SULFATE).
FT CARBOHYD 967 967 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1164 1164 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 1257 AA; 135544 MW; 992B33DCFA19EE1B CRC64;

Query Match 8.7%; Score 174; DB 1; Length 1257;
Best Local Similarity 30.8%; Pred. No. 1.8e-06;
Matches 44; Conservative 18; Mismatches 49; Indels 32; Gaps 6;

Qy 38 QRPCYKVIYFHDTSRRLNFEFEAKEACRRDGGQLVLSIESEDEQKLEKFIENLLPSDGDWF 97
Db 1037 QGHCVR--YF---AHRRAWEDARDRCRRRAGHLTSVHSPEEHKFINSF-----CHENSW 1085

Qy 98 IGLRRREEKQSNSTACQDLVATWTDGSIQFRNMYVYDEPS---CGSEVCVVMYHQPAPAG 154
Db 1086 IGLNDRTVRD-----FQTDNTGLQYENWREKQPDNFFAGGEDCVVMVAHENG--- 1134

Qy 155 IGGPYMFQWNDRCNMKNPFICK 177
Db 1135 -----RWNDVPCYNLNPYVCK 1150

RESULT 8
PGCV RAT
ID PGCV RAT STANDARD; PRT; 2738 AA.
AC Q9ERB4; O08592; O88564; Q9R1K4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Versican core protein precursor (large fibroblast proteoglycan)
DE (Chondroitin sulfate proteoglycan core protein 2) (PG-M) (Glial
DE hyaluronate-binding protein) (GHAP) (Fragments).
GN CPBG2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE OF 349-2738 FROM N.A. (ISOFORM V0), SEQUENCE FROM N.A.
RP (ISOFORM V3), AND SEQUENCE OF 2657-2738 FROM N.A. (ISOFORM VINT).
RC STRAIN=Wistar Kyoto;
RX MEDLINE=99327053; PubMed=10397680;
RA Lemire J.M., Braun K.R., Maurel P., Kaplan E.D., Schwartz S.M.,
RA Wight T.N.;
RT "Versican/PG-M isoforms in vascular smooth muscle cells.";
RL Arterioscler. Thromb. Vasc. Biol. 19:1630-1639(1999).
RN [2]
RP SEQUENCE OF 349-2738 FROM N.A. (ISOFORM V0).
RC STRAIN=Wistar Kyoto;
RX MEDLINE=98308094; PubMed=9642104;
RA Milev P., Maurel P., Chiba A., Mevissen M., Popp S., Yamaguchi Y.,

Db 2519 QGQYK--YF---AHRRTWDAARECRLOGAHLTSLSHEEQMFVNRV-----GHDIYQ 2566
QY 97 WIGLRRREKQSNSTACQDLYAWTDGSIQFRNWTYDEP---SCGSEVCVVMVHQPSAP 152
Db 2567 WIGL-----NDKMFHDFRWTGSLQYENWRPNQDPSFFSAGEDCVVIWHENG-- 2616
QY 153 AGIGGPFMFQWDDRCNMKNMFICKYS---DEKPAVPSREAEGE 193
Db 2617 -----QWNDVPCNVHLTYTCKGTGTVACGQPPVVENAKTFGK 2652

RESULT 9
PGCV_MOUSE
ID PGCV_MOUSE STANDARD; PRT; 3358 AA.
AC Q62059; Q62058; Q9CUU0;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Versican core protein precursor (Large fibroblast proteoglycan)
DE (Chondroitin sulfate proteoglycan core protein 2) (PG-M).
GN CSPG2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A. (ISOFORMS V0; V1 AND V2).
RC STRAIN=C57BL/6, and Swiss Webster; TISSUE=Brain;
RX MEDLINE=95122551; PubMed=7822336;
RA Ito K., Shinomura T., Zako M., Ujita M., Kimata K.;
RT "Multiple forms of mouse PG-M, a large chondroitin sulfate
RT proteoglycan generated by alternative splicing."
RL J. Biol. Chem. 270:958-965(1995).
RN [2]
SEQUENCE FROM N.A. (ISOFORM V3).
RP STRAIN=C57BL/6;
RC MEDLINE=95181355; PubMed=7876137;
RX Zako M., Shinomura T., Ujita M., Ito K., Kimata K.;
RA "Expression of PG-M(V3), an alternatively spliced form of PG-M
RT without a chondroitin sulfate attachment in region in mouse and human
RT tissues."
RL J. Biol. Chem. 270:3914-3918(1995).
RN [3]
SEQUENCE OF 1-1692 FROM N.A. (ISOFORM V1).
RC STRAIN=C57BL/6J; TISSUE=Skin;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombarts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
RN [4]
INTERACTION WITH FBLN1.
RX PubMed=10400671;
RA Aspberg A., Adam S., Kostka G., Timpl R., Heinegaard D.;
RT "Fibulin-1 is a ligand for the C-type lectin domains of aggrecan and
RT versican."

RL J. Biol. Chem. 274:20444-20449(1999).
CC -!- FUNCTION: May play a role in intercellular signaling and in
CC connecting cells with the extracellular matrix. May take part in
CC the regulation of cell motility, growth and differentiation. Binds
CC hyaluronic acid.
CC -!- SUBUNIT: Interacts with FBLN1.
CC -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=4;
CC Comment=Additional isoforms seem to exist;
CC Name=V0;
CC IsoId=Q62059-1; Sequence=Displayed;
CC Name=V1;
CC IsoId=Q62059-2; Sequence=VSP_003087, VSP_003088;
CC Name=V2;
CC IsoId=Q62059-3; Sequence=VSP_003089;
CC Name=V3;
CC IsoId=Q62059-4; Sequence=VSP_003087, VSP_003090;
CC -!- TISSUE SPECIFICITY: V2 is found only in brain.
CC -!- DEVELOPMENTAL STAGE: Disappears after the cartilage development.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -!- SIMILARITY: Contains 2 link domains.
CC -!- SIMILARITY: Contains 2 EGF-like domains.
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC -!- SIMILARITY: Contains 1 Sushi (SCR) domain.
CC -!- SIMILARITY: BELONGS TO THE AGGRECAN/VERSICAN PROTEOGLYCAN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D16263; BAA03796.1; --
DR EMBL; D28599; -; NOT ANNOTATED_CDS.
DR EMBL; D32040; BAA06802.1; --
DR EMBL; AK014525; BAB29411.1; --
DR HSSP; P01132; 1EPG.
DR MGD; MGI:102889; Cbpg2.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR007110; Ig-Like.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR000538; Link.
DR InterPro; IPR000436; Sushi_SCR_CCP.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00059; lectin_c; 1.
DR Pfam; PF00084; sushi; 1.
DR Pfam; PF00193; Xlink; 2.
DR PRINTS; PR01265; LINKMODULE.
DR ProDom; PD000918; Link; 2.
DR SMART; SM00032; CCP; 1.
DR SMART; SM00034; CLECT; 1.
DR SMART; SM00179; EGF_CA; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00445; LINK; 2.
DR PROSITE; PS00010; ASX_HYDROXYL; 1.
DR PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR PROSITE; PS00041; C_TYPE_LECTIN_2; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01187; EGF_CA; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR PROSITE; PS01241; LINK; 2.
KW Glycoprotein; Proteoglycan; Lectin; Extracellular matrix; Sushi;
Signal; Repeat; EGF-like domain; Calcium; Immunoglobulin domain;

QY 153 AGIGPYMFQWDDRCNMKNFNICKYS-----DEKPAVPSREAEGE 193
Db 3275 -----QWNDVPCNYHLTYTCKKGTACGQPPVVENAKTFGK 3310

RESULT 11
CD93_RAT
ID CD93_RAT STANDARD; PRT; 643 AA.
AC Q9ET61; Q9J126;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Complement component C1q receptor precursor (Complement component 1, q
DE subcomponent, receptor 1) (C1qRp) (C1qR(p)) (C1q/MBL/SPA receptor)
DE (CD93 antigen) (Cell surface antigen AA4).
GN C1QR1 OR CD93 OR C1QRP.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PVG; TISSUE=Natural killer cells;
RX MEDLINE=20545218; PubMed=11093152;
RA Lovik G., Vaage J.T., Dissen E., Szpirer C., Ryan J.C., Rolstad B.;
RT "Characterization and molecular cloning of rat C1qRp, a receptor on NK
cells.";
RL Eur. J. Immunol. 30:3355-3362(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar; TISSUE=Lung;
RX MEDLINE=20507883; PubMed=10934210;
RA Dean Y.D., McGreal E.P., Akatsu H., Gasque P.;
RT "Molecular and cellular properties of the rat AA4 antigen, a C-type
lectin-like receptor with structural homology to thrombomodulin.";
RL J. Biol. Chem. 275:34382-34392(2000).
CC -!- FUNCTION: Receptor (or element of a larger receptor complex) for
CC C1q, mannose-binding lectin (MBL2) and pulmonary surfactant
CC protein A (SPA). May mediate the enhancement of phagocytosis in
CC monocytes and macrophages upon interaction with soluble defense
CC collagens. May play a role in intercellular adhesion.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: Widely expressed. Highly expressed in lung and
CC heart. Expressed at lower level in brain, thymus, liver, spleen,
CC intestine, kidney, adrenal gland, muscle and testis. Expressed on
CC endothelial cells, platelets, undifferentiated monocytes and
CC circulating natural killer cells.
CC -!- PTM: N- and O-glycosylated (By similarity).
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC -!- SIMILARITY: Contains 5 EGF-like domains.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
or send an email to license@isb-sib.ch).

EMBL; AF136537; AAG01572.1; --
EMBL; AF160978; AAF80402.1; --
HSSP; P35555; 1EMN.
GO; GO:0016021; C:integral to membrane; ISS.
GO; GO:0004872; F:receptor activity; ISS.
GO; GO:0016337; P:cell-cell adhesion; ISS.
GO; GO:0042116; P:macrophage activation; ISS.
GO; GO:0006909; P:phagocytosis; ISS.
InterPro; IPR000152; Asx_hydroxyl.
InterPro; IPR001881; EGF_Ca.
InterPro; IPR006209; EGF_Like.
InterPro; IPR001304; Lectin_C.
Pfam; PF00008; EGF; 5.
Pfam; PF00059; lectin_c; 1.

DR SMART; SM00034; CLECT; 1.
DR SMART; SM00179; EGF CA; 3.
DR PROSITE; PS00010; ASX_HYDROXYL; 3.
DR PROSITE; PS00615; C_TYPE_LLECTIN_1; FALSE_NEG.
DR PROSITE; PS00041; C_TYPE_LLECTIN_2; 1.
DR PROSITE; PS01186; EGF_2; 3.
DR PROSITE; PS01187; EGF_CA; 3.
KW Cell adhesion; Receptor; Repeat; Signal; Transmembrane;
KW EGF-like domain; Lectin; Glycoprotein.
FT SIGNAL 1 23
FT CHAIN 24 643
FT DOMAIN 24 571
FT TRANSMEM 572 592
FT DOMAIN 593 643
FT DOMAIN 31 173
FT DOMAIN 257 298
FT DOMAIN 299 341
FT DOMAIN 342 381
FT DOMAIN 382 423
FT DOMAIN 424 462
FT DISULFID 261 272
FT DISULFID 268 282
FT DISULFID 284 297
FT DISULFID 303 314
FT DISULFID 308 325
FT DISULFID 327 340
FT DISULFID 346 355
FT DISULFID 351 364
FT DISULFID 366 380
FT DISULFID 386 397
FT DISULFID 393 406
FT DISULFID 408 422
FT DISULFID 428 437
FT DISULFID 433 446
FT DISULFID 448 461
FT CARBOHYD 322 322
FT CARBOHYD 498 498
FT CONFLICT 417 417
SQ SEQUENCE 643 AA; 68781 MW; 9AE4C933AD943DB6 CRC64;
Query Match 8.6%; Score 172.5; DB 1; Length 643;
Best Local Similarity 25.9%; Pred. No. 9.8e-07;
Matches 56; Conservative 37; Mismatches 74; Indels 49; Gaps 11;
QY 14 VLLVGLRAATGRLLSGP-----VCRGGTQPCYKVIYFHDTSRRLNFEEAKEACRRDG 67
Db LLLLGL---LGQWAGAAADSEAVVCEG---TACYTAHW-----GKLSAAEAQRCNENG 56
QY 68 GQLVSISEDEQKLEKFIENLL---PSD---GDFWIGLRRREEKQSNSTACQDLYAWT 120
Db GNLATVKSSEEARHVQEAALQLLTKAPSETKIGKFWIGLQREKGCCTYHDLPMKGFSWV 116
QY 121 -DGSISQFRNWY-VDEPSCGSEVCVWY-----HQSAPAGIGGPPYMFQWDDRC--- 168
Db GGEDTTSYNNWKASKSSCISKRCVSLILDLSLKPSPHLP-----KWHSPCGTP 167
QY 169 ----NMKNFNICKYSDEKPAVPSREAEGETELTTP 200
Db DAPGNSIEGFLCKFNFKMGCSPLALGGPGQLTYTTP 203

RESULT 12
PGCV_CHICK
ID PGCV_CHICK STANDARD; PRT; 3562 AA.
AC Q90953; Q90945;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Versican core protein precursor (Large fibroblast proteoglycan)
DE (Chondroitin sulfate proteoglycan core protein 2) (PG-M).
GN CPBG2.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS V0 AND V1).
RC STRAIN=White leghorn; TISSUE=Limb bud;
RX MEDLINE=93300846; PubMed=8314802;
RA Shinomura T., Nishida Y., Ito K., Kimata K.;
RT "cDNA cloning of PG-M, a large chondroitin sulfate proteoglycan
RT expressed during chondrogenesis in chick limb buds. Alternative
RT spliced multiforms of PG-M and their relationships to versican.";
RL J. Biol. Chem. 268:14461-14469(1993).
CC -!- FUNCTION: May play a role in intercellular signaling and in
CC connecting cells with the extracellular matrix. May take part in
CC the regulation of cell motility, growth and differentiation. Binds
CC hyaluronic acid.
CC -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Comment=Additional isoforms seem to exist;
CC Name=V0;
CC IsoId=Q090953-1; Sequence=Displayed;
CC Name=V1;
CC IsoId=Q090953-2; Sequence=VSP_003093;
CC TISSUE SPECIFICITY: Prechondrogenic condensation area of
CC developing limb buds.
CC -!- DEVELOPMENTAL STAGE: Disappears after the cartilage development
CC (By similarity).
CC -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -!- SIMILARITY: Contains 2 link domains.
CC -!- SIMILARITY: Contains 2 EGF-like domains.
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC -!- SIMILARITY: Contains 1 Sushi (SCR) domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X60226; CAA42787.1; -;
DR EMBL; D13542; BAA02742.1; -;
DR FIR; A47171; A47171.
DR HSSP; P00740; 1EDM.
DR InterPro; IPR000152; Asx hydroxyl.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR000538; Link.
DR InterPro; IPR000436; Sushi_SCR_CCP.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF00059; lectin_c; 1.
DR Pfam; PF00084; sushi; 1.
DR Pfam; PF00193; Xlink; 2.
DR PRINTS; PR01265; LINKMODULE.
DR PRODOM; PD000918; Link; 2.
DR SMART; SM00032; CCP; 1.
DR SMART; SM00034; CLECT; 1.
DR SMART; SM00179; EGF_CA; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00445; LINK; 2.
DR PROSITE; PS00010; ASX_HYDROXYL; 1.
DR PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.

DR PROSITE; PS01187; EGF_CA; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR PROSITE; PS01241; LINK; 2.
KW Glycoprotein; Proteoglycan; Lectin; Extracellular matrix; Sushi;
KW Signal; Repeat; EGF-like domain; Calcium; Immunoglobulin domain;
KW Hyaluronic acid; Alternative splicing.
FT SIGNAL 1 26
FT CHAIN 27 3562 VERSICAN CORE PROTEIN.
FT DOMAIN 27 143 IG-LIKE V-TYPE.
FT DOMAIN 166 243 LINK 1.
FT DOMAIN 264 345 LINK 2.
FT DOMAIN 3254 3290 EGF-LIKE 1.
FT DOMAIN 3292 3328 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 3341 3455 C-TYPE LECTIN.
FT DOMAIN 3460 3518 SUSHI.
FT DISULFID 44 129 BY SIMILARITY.
FT DISULFID 171 242 BY SIMILARITY.
FT DISULFID 195 216 BY SIMILARITY.
FT DISULFID 269 344 BY SIMILARITY.
FT DISULFID 293 314 BY SIMILARITY.
FT DISULFID 3258 3269 BY SIMILARITY.
FT DISULFID 3263 3278 BY SIMILARITY.
FT DISULFID 3280 3289 BY SIMILARITY.
FT DISULFID 3296 3307 BY SIMILARITY.
FT DISULFID 3301 3316 BY SIMILARITY.
FT DISULFID 3318 3327 BY SIMILARITY.
FT DISULFID 3334 3345 BY SIMILARITY.
FT DISULFID 3362 3454 BY SIMILARITY.
FT DISULFID 3430 3446 BY SIMILARITY.
FT DISULFID 3461 3504 BY SIMILARITY.
FT DISULFID 3490 3517 BY SIMILARITY.
FT CARBOHYD 163 163 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 235 235 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 329 329 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 529 529 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 709 709 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 948 948 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1409 1409 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1479 1479 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1523 1523 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1530 1530 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1625 1625 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1751 1751 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1988 1988 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2088 2088 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2089 2089 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2507 2507 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2642 2642 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2679 2679 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2748 2748 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2762 2762 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 3069 3069 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 3194 3194 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 3232 3232 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 3545 3545 N-LINKED (GLCNAC. .) (POTENTIAL).
FT VARSPLIC 485 1411 Missing (in isoform V1).
FT SQ SEQUENCE 3562 AA; 388078 MW; 9BC566E88C1602D2 CRC64;
/FTid=VSP_003093.

Query Match 8.6%; Score 171; DB 1; Length 3562;
Best Local Similarity 28.5%; Pred. No. 1.3e-05;
Matches 47; Conservative 23; Mismatches 55; Indels 40; Gaps 8;
QY 38 QRPCYKVIYFHDTSRRRLNFEAEKACRRDGGQLVSISEDEQKLEFIENLLPSDGF- 96
Db 3342 QGQCYK--YF---AHRRTWDTAERECRLQGAHLTSLSHEEQVFVNRI-----GHDYQ 3389

QY 97 WIGLRRREKQSNSTACQDLVATWDGSIQFRNYYVDEP----SCGSEVCVVMYHQPSAP 152
Db 3390 WIGL-----NDKMFERDFRWTGSPLOQYENWRPNQDPSFSGEDCVIWHENG-- 3439
QY 153 AGIGGPFYMFQWNDRCNMKNFICKYS----DEKPAVPSREAEGE 193

Db 3440 -----QMNVDPCNYHLTYTCKKGTVACGQPPVVENAKTFGK 3475

RESULT 13

PGCA BOVIN

ID PGCA BOVIN STANDARD; PRT; 2364 AA.

AC P13608; P79117; Q28159;

DT 01-JAN-1990 (Rel. 13, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Aggrecan core protein precursor (Cartilage-specific proteoglycan core

DE protein) (CSPCP).

GN AGC1.

OS Bos taurus (Bovine).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

OC Bovidae; Bovinae; Bos.

OX NCBI_TaxID=9913;

RN [1]

RP SEQUENCE FROM N.A.

RA Hering T.M., Kollar J., Huynh T.D.;

RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE OF 563-1056 FROM N.A.

RX MEDLINE=89380219; PubMed=2528543;

RA Antonsson P., Heinegaard D., Oldberg A.;

RT "The keratan sulfate-enriched region of bovine cartilage proteoglycan

RT consists of a consecutively repeated hexapeptide motif.";

RL J. Biol. Chem. 264:16170-16173(1989).

RN [3]

RP SEQUENCE OF 1609-2113 AND 2151-2364 FROM N.A.

RX MEDLINE=87270630; PubMed=3111460;

RA Oldberg A., Antonsson P., Heinegaard D.;

RT "The partial amino acid sequence of bovine cartilage proteoglycan,

RT deduced from a cDNA clone, contains numerous Ser-Gly sequences

RT arranged in homologous repeats.";

RL Biochem. J. 243:255-259(1987).

RN [4]

RP SEQUENCE OF 2114-2150 FROM N.A.

RC TISSUE=Cartilage;

RX MEDLINE=93352525; PubMed=8349621;

RA Fueloep C., Walcz E., Valyon M., Glant T.T.;

RT "Expression of alternatively spliced epidermal growth factor-like

RT domains in aggrecans of different species. Evidence for a novel

RT module.";

RL J. Biol. Chem. 268:17377-17383(1993).

RN [5]

RP PARTIAL SEQUENCE.

RX MEDLINE=85027710; PubMed=6489519;

RA Perin J.-P., Bonnet F., Jolles J., Jolles P.;

RT "Sequence data concerning the protein core of the cartilage

RT proteoglycan monomers. Characterization of a sequence allowing the

RT synthesis of an oligonucleotide probe.";

RL FEBS Lett. 176:37-42(1984).

RN [6]

RP PARTIAL SEQUENCE.

RX MEDLINE=87005253; PubMed=3530809;

RA Perin J.P., Bonnet F., Jolles P.;

RT "Structural relationship between link proteins and proteoglycan

RT monomers.";

RL FEBS Lett. 206:73-77(1986).

CC -!- FUNCTION: THIS PROTEOGLYCAN IS A MAJOR COMPONENT OF EXTRACELLULAR

CC MATRIX OF CARTILAGENOUS TISSUES. A MAJOR FUNCTION OF THIS PROTEIN

CC IS TO RESIST COMPRESSION IN CARTILAGE. IT BINDS AVIDLY TO

CC HYALURONIC ACID VIA AN AMINO-TERMINAL GLOBULAR REGION. MAY PLAY A

CC REGULATORY ROLE IN THE MATRIX ASSEMBLY OF THE CARTILAGE.

CC -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix (By

CC similarity).

CC -!- ALTERNATIVE PRODUCTS:

CC Event=Alternative splicing; Named isoforms=2;

CC Name=1;

CC IsoId=P13608-1; Sequence=Displayed;

CC Name=2;

CC

CC IsoId=P13608-2; Sequence=VSP_003072;

CC -!- DOMAIN: TWO GLOBULAR DOMAINS, G1 AND G2, COMPRISE THE AMINO

CC TERMINUS OF THE PROTEOGLYCAN, WHILE ANOTHER GLOBULAR REGION, G3,

CC MAKES UP THE C-TERMINUS. G1 CONTAINS LINK DOMAINS AND THUS

CC CONSISTS OF THREE DISULFIDE-BONDED LOOP STRUCTURES DESIGNATED AS

CC THE A, B, B' MOTIFS. G2 IS SIMILAR TO G1. THE KERATAN SULFATE (KS)

CC AND THE CHONDROITIN SULFATE (CS) ATTACHMENT DOMAINS LIE BETWEEN G2

CC AND G3.

CC -!- PTM: CONTAINS MOSTLY CHONDROITIN SULFATE, BUT ALSO N-LINKED AND O-

CC LINKED (ABOUT 40) OLIGOSACCHARIDES.

CC -!- PTM: THE KERATAN SULFATE CONTENTS DIFFER CONSIDERABLY BETWEEN

CC ADULT AND FETAL BOVINE PROTEOGLYCANS.

CC -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.

CC -!- SIMILARITY: Contains 4 link domains.

CC -!- SIMILARITY: Contains 1 EGF-like domain.

CC -!- SIMILARITY: Contains 1 C-type lectin family domain.

CC -!- SIMILARITY: Contains 1 Sushi (SCR) domain.

CC -!- SIMILARITY: BELONGS TO THE AGGREGAN/VERSICAN PROTEOGLYCAN FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch>) or

CC send an email to license@isb-sib.ch.

CC -----

CC EMBL; U76615; AAB38524.1; -.

CC EMBL; L07053; -; NOT_ANNOTATED_CDS.

CC PIR; A34234; A39808.

CC PIR; T42630; T42630.

CC HSSP; P08709; 1BF9.

CC InterPro; IPR002353; AntifreezeII.

CC InterPro; IPR000152; Asx_hydroxyl.

CC InterPro; IPR000742; EGF_2.

CC InterPro; IPR001881; EGF_Ca.

CC InterPro; IPR006209; EGF_like.

CC InterPro; IPR007110; Ig_Like.

CC InterPro; IPR003006; Ig_MHC.

CC InterPro; IPR001304; Lectin_C.

CC InterPro; IPR000538; Link.

CC InterPro; IPR003324; SGXXSG.

CC InterPro; IPR000436; Sushi_SCR_CCP.

CC Pfam; PF00008; EGF; 1.

CC Pfam; PF00047; Ig; 1.

CC Pfam; PF00059; lectin_c; 1.

CC Pfam; PF02339; SGXXSG; 61.

CC Pfam; PF00084; sushi; 1.

CC Pfam; PF00193; Xlink; 4.

CC PRINTS; PR01265; LINKMODULE.

CC PRINTS; PR00356; ANTIFREEZEII.

CC ProDom; PD000918; Link; 4.

CC SMART; SM00032; CCP; 1.

CC SMART; SM00034; CLECT; 1.

CC SMART; SM00179; EGF_CA; 1.

CC SMART; SM00445; LINK; 4.

CC PROSITE; PS00010; ASX_HYDROXYL; 1.

CC PROSITE; PS00615; C_TYPE_LLECTIN_1; 1.

CC PROSITE; PS50041; C_TYPE_LLECTIN_2; 1.

CC PROSITE; PS00022; EGF_1; 1.

CC PROSITE; PS01187; EGF_CA; 1.

CC PROSITE; PS50835; IG_LIKE; 1.

CC PROSITE; PS00290; IG_MHC; FALSE_NEG.

CC PROSITE; PS01241; LINK; 4.

CC KW Glycoprotein; Cartilage; Proteoglycan; Lectin; Signal; Sushi;

CC KW EGF-like domain; Calcium; Alternative splicing; Repeat;

CC KW Immunoglobulin domain.

CC SIGNAL 1 16 POTENTIAL.

CC CHAIN 17 2364 AGGREGAN CORE PROTEIN.

CC DOMAIN 25 147 IG-LIKE V-TYPE.

CC DOMAIN 170 247 LINK 1.

CC DOMAIN 268 349 LINK 2.

CC DOMAIN 504 581 LINK 3.

FT	DOMAIN	602	683	LINK 4.	
FT	DOMAIN	774	907	23 X 6 AA APPROXIMATE TANDEM REPEATS OF E-[EK]-P-F-P-S.	
FT	DOMAIN	1433	2112	CS-2.	
FT	DOMAIN	2113	2149	EGF-LIKE, CALCIUM-BINDING (POTENTIAL).	
FT	DOMAIN	2114	2364	G3.	
FT	DOMAIN	2161	2276	C-TYPE LECTIN.	
FT	DOMAIN	2280	2338	SUSHI.	
FT	DISULFID	51	133	BY SIMILARITY.	
FT	DISULFID	175	246	BY SIMILARITY.	
FT	DISULFID	199	220	BY SIMILARITY.	
FT	DISULFID	273	348	BY SIMILARITY.	
FT	DISULFID	297	318	BY SIMILARITY.	
FT	DISULFID	509	580	BY SIMILARITY.	
FT	DISULFID	533	554	BY SIMILARITY.	
FT	DISULFID	607	682	BY SIMILARITY.	
FT	DISULFID	631	652	BY SIMILARITY.	
FT	DISULFID	2117	2128	BY SIMILARITY.	
FT	DISULFID	2182	2274	BY SIMILARITY.	
FT	DISULFID	2250	2266	BY SIMILARITY.	
FT	DISULFID	2281	2324	BY SIMILARITY.	
FT	DISULFID	2310	2337	BY SIMILARITY.	
FT	CARBOHYD	126	126	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	239	239	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	333	333	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	387	387	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	611	611	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	667	667	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	VARSPLIC	2114	2150	Missing (in isoform 2).	
FT	SEQUENCE	2364 AA;	246359 MW;	6FF83763420C3D4C CRC64;	
SQ	SEQUENCE	2364 AA;	246359 MW;	6FF83763420C3D4C CRC64;	
Query Match					7.9%; Score 158.5; DB 1; Length 2364;
Best Local Similarity					26.8%; Pred. No. 8.6e-05;
Matches					48; Conservative 29; Mismatches 59; Indels 43; Gaps 10;
QY	30 QPVCRRGG---	TQRPCYKVIYFHDTSRRLNFEFEAKEACRRDGGQLVSI	ESEDEQKLIKFI	86	
Db	2151 QKLCEEGWTKFQGHCHYR--	HFDP---RATWVDAESQCRKQKQSHLSSIVTPEEQ---	EFV	2201	
QY	87 ENLLPSDGF-	WIGLRRREKQSNSTACQDLYAWTDGSI	SQFRNYYVDEP----	SCGSEV	141
Db	2202 NN---	NAQDYQWIGL-----NDKTIEGDFRWSGDGHS	LQFENWRPNQPNPFATGEDC	2250	
QY	142 CVVMYHQPSAPAGIGGYPMFQWDDRCNMKNFICKYS---	DEKPAVPSREAEGETE	196		
Db	2251 VVMIWHEKG-----	EWNDVPCNYQLPFTCKGTACGEPVVEHARIFGQKKD	2298		
RESULT 14					
PGCA_HUMAN					
ID	PGCA_HUMAN	STANDARD;	PRT;	2415 AA.	
AC	P16112; Q13650;				
DT	01-APR-1990 (Rel. 14, Created)				
DT	01-AUG-1992 (Rel. 23, Last sequence update)				
DT	15-SEP-2003 (Rel. 42, Last annotation update)				
DE	Aggrecan core protein precursor (Cartilage-specific proteoglycan core protein)				
DE	(CSPCP) (Chondroitin sulfate proteoglycan core protein 1).				
GN	AGC1 OR CSPG1.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A. (ISOFORM 3).				
RC	TISSUE=Chondrocytes;				
RX	MEDLINE=91093289; PubMed=1985970;				
RA	Dooge K.J., Sasaki M., Kimura T., Yamada Y.;				
RT	"Complete coding sequence and deduced primary structure of the human				
RT	cartilage large aggregating proteoglycan, aggrecan. Human-specific				
RT	repeats, and additional alternatively spliced forms."				
RL	J. Biol. Chem. 266:894-902(1991).				
RN	[2]				

RP	SEQUENCE OF 1778-2415 FROM N.A. (ISOFORM 2).	
RC	TISSUE=Chondrocytes;	
RA	Dudhia J., Hardingham T.E.;	
RL	Submitted (JAN-1990) to the EMBL/GenBank/DBJ databases.	
RN	[3]	
RP	SEQUENCE OF 1936-2415 FROM N.A. (ISOFORM 1).	
RX	MEDLINE=89380154; PubMed=2789216;	
RA	Baldwin C.T., Reginato A.M., Prockop D.J.;	
RT	"A new epidermal growth factor-like domain in the human core protein	
RT	for the large cartilage-specific proteoglycan. Evidence for	
RT	alternative splicing of the domain."	
RL	J. Biol. Chem. 264:15747-15750(1989).	
RN	[4]	
RP	SEQUENCE OF 764-864 FROM N.A.	
RC	TISSUE=Blood;	
RX	MEDLINE=95128522; PubMed=7827755;	
RA	Barry F.P., Neame P.J., Sasse J., Pearson D.;	
RT	"Length variation in the keratan sulfate domain of mammalian	
RT	aggrecan."	
RL	Matrix Biol. 14:323-328(1994).	
CC	-1- FUNCTION: THIS PROTEOGLYCAN IS A MAJOR COMPONENT OF EXTRACELLULAR	
CC	MATRIX OF CARTILAGINOUS TISSUES. A MAJOR FUNCTION OF THIS PROTEIN	
CC	IS TO RESIST COMPRESSION IN CARTILAGE. IT BINDS AVIDLY TO	
CC	HYALURONIC ACID VIA AN AMINO-TERMINAL GLOBULAR REGION.	
CC	-1- SUBUNIT: Interacts with FBLN1 (By similarity).	
CC	-1- SUBCELLULAR LOCATION: Secreted; extracellular matrix (By	
CC	similarity).	
CC	-1- ALTERNATIVE PRODUCTS:	
CC	Event=Alternative splicing; Named isoforms=3;	
CC	Comment=Additional isoforms seem to exist;	
CC	Name=1;	
CC	ISOId=P161112-1; Sequence=Displayed;	
CC	Name=2;	
CC	ISOId=P161112-2; Sequence=VSP_003074;	
CC	Name=3;	
CC	ISOId=P161112-3; Sequence=VSP_003074, VSP_003075;	
CC	-1- DOMAIN: TWO GLOBULAR DOMAINS, G1 AND G2, COMPRISE THE AMINO	
CC	TERMINUS OF THE PROTEOGLYCAN, WHILE ANOTHER GLOBULAR REGION, G3,	
CC	MAKES UP THE COOH TERMINUS. G1 CONTAINS LINK DOMAINS AND THUS	
CC	CONSISTS OF THREE DISULFIDE-BONDED LOOP STRUCTURES DESIGNATED AS	
CC	THE A, B, B' MOTIFS. G2 IS SIMILAR TO G1. THE KERATAN SULFATE (KS)	
CC	AND THE CHONDROITIN SULFATE (CS) ATTACHMENT DOMAINS LIE BETWEEN G2	
CC	AND G3.	
CC	-1- PTM: CONTAINS MOSTLY CHONDROITIN SULFATE, BUT ALSO KERATAN SULFATE	
CC	CHAINS, N-LINKED AND O-LINKED OLIGOSACCHARIDES.	
CC	-1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.	
CC	-1- SIMILARITY: Contains 4 link domains.	
CC	-1- SIMILARITY: Contains 1 EGF-like domain.	
CC	-1- SIMILARITY: Contains 1 C-type lectin family domain.	
CC	-1- SIMILARITY: Contains 1 Sushi (SCR) domain.	
CC	-1- SIMILARITY: BELONGS TO THE AGGRECAN/VERSICAN PROTEOGLYCAN FAMILY.	
CC	-----	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -	
CC	the European Bioinformatics Institute. There are no restrictions on its	
CC	use by non-profit institutions as long as its content is in no way	
CC	modified and this statement is not removed. Usage by and for commercial	
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/	
CC	or send an email to license@isb-sib.ch).	
CC	-----	
DR	EMBL; M55172; AAA62824.1; -	
DR	EMBL; J05062; AAA35726.1; -	
DR	EMBL; X17406; CAA35463.1; -	
DR	EMBL; S74659; AAC60643.2; -	
DR	PIR; A39086; A39086.	
DR	HSSP; P98066; 1TSG.	
DR	Genew; HGNC:319; AGC1.	
DR	MIM; 155760; -	
DR	GO; GO:0005204; F.chondroitin sulfate proteoglycan; TAS.	
DR	InterPro; IPR002353; Antifreeze1.	
DR	InterPro; IPR006209; EGF-like.	
DR	InterPro; IPR006210; IEGF.	
DR	InterPro; IPR007110; Ig-like.	

DR InterPro; IPR003599; Ig.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR000538; Link.
DR InterPro; IPR003324; SGXXSG.
DR InterPro; IPR000436; Sushi_SCR_CCP.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF00059; lectin_c; 1.
DR Pfam; PF02339; SGXXSG; 71.
DR Pfam; PF00084; sushi; 1.
DR Pfam; PF00193; Xlink; 4.
DR PRINTS; PR01265; LINKMODULE.
DR PRINTS; PR00356; ANTIFREEZEII.
DR ProDom; PD000918; Link; 4.
DR SMART; SM00032; CCP; 1.
DR SMART; SM00034; CLECT; 1.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00445; LINK; 4.
DR PROSITE; PS00615; C-TYPE LECTIN_1; 1.
DR PROSITE; PS00041; C-TYPE LECTIN_2; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00835; IG_LIKE; 1.
DR PROSITE; PS00290; IG_MHC; 1.
DR PROSITE; PS01241; LINK; 3.
KW Glycoprotein; Carilage; Proteoglycan; Lectin; Signal; Sushi;
KW EGF-like domain; Alternative splicing; Repeat; Immunoglobulin domain.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 2415 AGGREGAN CORE PROTEIN.
FT DOMAIN 34 147 IG-LIKE V-TYPE.
FT DOMAIN 170 247 LINK 1.
FT DOMAIN 268 349 LINK 2.
FT DOMAIN 495 572 LINK 3.
FT DOMAIN 593 673 LINK 4.
FT DOMAIN 2164 2199 EGF-LIKE.
FT DOMAIN 2201 2327 C-TYPE LECTIN.
FT DOMAIN 2331 2389 SUSHI.
FT DOMAIN 48 141 G1-A.
FT DOMAIN 152 247 G1-B.
FT DOMAIN 253 349 G1-B'.
FT DOMAIN 477 571 G2-B.
FT DOMAIN 578 672 G2-B'.
FT DOMAIN 676 848 KS.
FT DOMAIN 772 844 12 X APPROXIMATE TANDEM REPEATS.
FT DOMAIN 851 1497 CS-1.
FT DOMAIN 941 1497 29 X APPROXIMATE TANDEM REPEATS.
FT DOMAIN 1498 2162 CS-2.
FT DOMAIN 2163 2415 G3.
FT DISULFID 51 133 BY SIMILARITY.
FT DISULFID 175 246 BY SIMILARITY.
FT DISULFID 199 220 BY SIMILARITY.
FT DISULFID 273 348 BY SIMILARITY.
FT DISULFID 297 318 BY SIMILARITY.
FT DISULFID 500 571 BY SIMILARITY.
FT DISULFID 524 545 BY SIMILARITY.
FT DISULFID 598 672 BY SIMILARITY.
FT DISULFID 621 642 BY SIMILARITY.
FT DISULFID 2168 2178 BY SIMILARITY.
FT DISULFID 2173 2187 BY SIMILARITY.
FT DISULFID 2189 2198 BY SIMILARITY.
FT DISULFID 2205 2216 BY SIMILARITY.
FT DISULFID 2233 2325 BY SIMILARITY.
FT DISULFID 2301 2317 BY SIMILARITY.
FT DISULFID 2332 2375 BY SIMILARITY.
FT DISULFID 2361 2388 BY SIMILARITY.
FT CARBOHYD 126 126 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 239 239 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 333 333 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 387 387 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 434 434 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 602 602 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 657 657 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 737 737 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1898 1898 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 2163 2200 Missing (in isoform 2 and isoform 3).
FT VARSPLIC 2330 2390 /FTid=VSP_003074.
FT CONFLICT 766 766 /FTid=VSP_003075.
FT CONFLICT 847 847 E -> A (IN REF. 4).
FT CONFLICT 1928 1928 E -> V (IN REF. 4).
FT CONFLICT 1964 1964 I -> V (IN REF. 2).
FT CONFLICT 2070 2070 P -> A (IN REF. 2 AND 3).
FT CONFLICT 2391 2391 A -> P (IN REF. 2 AND 3).
SQ SEQUENCE 2415 AA; 250191 MW; 1288937E1B98C6B6 CRC64;
Query Match 7.9%; Score 158.5; DB 1; Length 2415;
Best Local Similarity 29.1%; Pred. No. 8.9e-05;
Matches 52; Conservative 22; Mismatches 62; Indels 43; Gaps 11;
QY 30 QPVCRGG---TQPCYKVIYFHDTSRLNFEAEKACRRDGGQLVSEDEQKLEKFI 86
Db 2202 QEVEEGWKNYQGHYR--HFPD---RETWVDAERRCREQQSHLSSIVTPEEQ----EFV 2252
QY 87 ENLLPSDGDGFWGLRRREKQSNSTACQDLYAWTDGSGISQFRNMYVDEPS---CGSEVC 142
Db 2253 NN---NAQDYQWIGL-----NDRTEIGDFRWSGDGHPMQFENWRNPQDNFFAAGDC 2301
QY 143 VVM-YHQPSAPAGIGGPMFQWDDRCNMKNVFIKYS----DEKPAVPSREAEGETE 196
Db 2302 VVMWHEKG-----EWNDVPCNYHLPTCKKGTVACGEPVVEHARTFGQKGD 2349
RESULT 15
CD93 HUMAN
ID CD93 HUMAN STANDARD; PRT; 652 AA.
AC Q9NPY3; O00274;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Complement component Clq receptor precursor (Complement component 1, q
DE subcomponent, receptor 1) (ClqRp) (ClqR(p)) (Clq/MBL/SPA receptor)
DE (CD93 antigen) (CDw93).
GN ClQR1 OR CD93.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=97199258; PubMed=9047234;
RA Nepomuceno R.R., Henschen-Edman A.H., Burgess W.H., Tenner A.J.;
RT "CDNA cloning and primary structure analysis of ClqR(p), the human
RT Clq/MBL/SPA receptor that mediates enhanced phagocytosis in vitro.";
RL Immunity 6:119-129(1997).
RN [2]
RP SEQUENCE FROM N.A., AND VARIANT ALA-318.
RX MEDLINE=21640567; PubMed=11781389;
RA Steinberger P., Szekeres A., Wille S., Stockl J., Selenko N.,
RA Prager E., Staffler G., Madic O., Stockinger H., Knapp W.;
RT "Identification of human CD93 as the phagocytic Clq receptor (ClqRp)
RT by expression cloning.";
RL J. Leukoc. Biol. 71:133-140(2002).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=21638749; PubMed=11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,

Search completed: December 22, 2003, 16:10:57
Job time : 13.9128 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 22, 2003, 16:06:34 ; Search time 18.1827 Seconds
(without alignments)
1978.090 Million cell updates/sec

Title: US-09-887-855-2
Perfect score: 2000
Sequence: 1 MRPGTALQAVLLAVLLVGLR.....PDQMGRSKESGWENEIYGY 374

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	183	9.2	1456	1 A36563	mannose receptor p
2	178.5	8.9	1455	1 A48925	mannose receptor p
3	177	8.8	1268	2 S52781	neurocan - mouse
4	174.5	8.7	1643	2 T14274	versican precursor
5	174.5	8.7	3381	2 T42389	versican precursor
6	174	8.7	1257	2 S28764	neurocan precursor
7	174	8.7	2397	1 A55535	versican precursor
8	174	8.7	2409	1 A60979	versican precursor
9	171	8.6	3562	2 A47171	chondroitin sulfat
10	170.5	8.5	1479	2 T42710	mannose receptor,
11	158.5	7.9	1340	2 A39808	proteoglycan core
12	158.5	7.9	2327	2 T42630	aggrecan - bovine
13	158.5	7.9	2415	1 A39086	aggrecan precursor
14	156	7.8	162	1 LNRC1	lectin BRA3-1 prec
15	154.5	7.7	612	2 B42755	E-selectin precurs
16	153.5	7.7	2124	2 A28452	proteoglycan core
17	152	7.6	912	2 A54423	brevican precursor
18	151	7.5	162	1 LNRC3	lectin BRA3-2 prec
19	150.5	7.5	459	2 T24425	hypothetical prote
20	150	7.5	330	2 T46256	brevican - human (
21	149	7.4	321	1 LNHUER	IgE Fc receptor II
22	148.5	7.4	152	2 JC4690	coagulation factor
23	148.5	7.4	404	2 A46274	HIV gp120-binding
24	148.5	7.4	2132	1 A55182	aggrecan precursor
25	148	7.4	253	2 E89130	proteoglycan core
26	147	7.3	883	2 S57653	brevican precursor
27	146	7.3	2109	1 I50421	aggrecan precursor
28	145.5	7.3	742	2 JC7595	scavenger receptor
29	145	7.2	883	2 S49126	brevican precursor

30	144.5	7.2	173	2 S10548	lectin - barnacle
31	144.5	7.2	372	2 S23936	L-selectin precurs
32	143.5	7.2	129	2 JC4329	coagulation factor
33	143.5	7.2	372	1 A32375	L-selectin precurs
34	143.5	7.2	463	2 T26655	hypothetical prote
35	143	7.1	321	2 T26152	hypothetical prote
36	142.5	7.1	131	2 JC5058	bitiscetin alpha c
37	142	7.1	1487	2 S48719	phospholipase-A(2)
38	141.5	7.1	165	2 A47148	reg I, regeneratin
39	141.5	7.1	331	1 LNMSE	IgE Fc receptor, l
40	140	7.0	370	2 S22124	L-selectin precurs
41	139	7.0	248	1 LNHUPS	pulmonary surfacta
42	139	7.0	248	1 LNHUP6	pulmonary surfacta
43	139	7.0	248	1 LNHUP1	pulmonary surfacta
44	139	7.0	283	1 LNFHLS	lectin precursor -
45	138.5	6.9	172	2 S32489	lectin - Iberian r

ALIGNMENTS

RESULT 1

A36563

mannose receptor precursor - human

C;Species: Homo sapiens (man)

C;Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text change 10-Sep-1999

C;Accession: A36563; A60926; A44255; B44255; C44255; D44255; E44255; F44255; G44255; H44255; I44255; J44255; K44255; L44255; M44255; N44255; O44255; P44255; Q44255; R44255; S44255; T44255; U44255; V44255; W44255; X44255; Y44255; Z44255; AA44255; AB44255; AC44255; AD44255; AE44255; AF44255; AG44255; AH44255; AI44255; AJ44255; AK44255; AL44255; AM44255; AN44255; AO44255; AP44255; AQ44255; AR44255; AS44255; AT44255; AU44255; AV44255; AW44255; AX44255; AY44255; AZ44255; BA44255; BB44255; BC44255; BD44255; BE44255; BF44255; BG44255; BH44255; BI44255; BJ44255; BK44255; BL44255; BM44255; BN44255; BO44255; BP44255; BQ44255; BR44255; BS44255; BT44255; BU44255; BV44255; BW44255; BX44255; BY44255; BZ44255; CA44255; CB44255; CC44255; CD44255; CE44255; CF44255; CG44255; CH44255; CI44255; CJ44255; CK44255; CL44255; CM44255; CN44255; CO44255; CP44255; CQ44255; CR44255; CS44255; CT44255; CU44255; CV44255; CW44255; CX44255; CY44255; CZ44255; DA44255; DB44255; DC44255; DD44255; DE44255; DF44255; DG44255; DH44255; DI44255; DJ44255; DK44255; DL44255; DM44255; DN44255; DO44255; DP44255; DQ44255; DR44255; DS44255; DT44255; DU44255; DV44255; DW44255; DX44255; DY44255; DZ44255; EA44255; EB44255; EC44255; ED44255; EE44255; EF44255; EG44255; EH44255; EI44255; EJ44255; EK44255; EL44255; EM44255; EN44255; EO44255; EP44255; EQ44255; ER44255; ES44255; ET44255; EU44255; EV44255; EW44255; EX44255; EY44255; EZ44255; FA44255; FB44255; FC44255; FD44255; FE44255; FF44255; FG44255; FH44255; FI44255; FJ44255; FK44255; FL44255; FM44255; FN44255; FO44255; FP44255; FQ44255; FR44255; FS44255; FT44255; FU44255; FV44255; FW44255; FX44255; FY44255; FZ44255; GA44255; GB44255; GC44255; GD44255; GE44255; GF44255; GG44255; GH44255; GI44255; GJ44255; GK44255; GL44255; GM44255; GN44255; GO44255; GP44255; GQ44255; GR44255; GS44255; GT44255; GU44255; GV44255; GW44255; GX44255; GY44255; GZ44255; HA44255; HB44255; HC44255; HD44255; HE44255; HF44255; HG44255; HH44255; HI44255; HJ44255; HK44255; HL44255; HM44255; HN44255; HO44255; HP44255; HQ44255; HR44255; HS44255; HT44255; HU44255; HV44255; HW44255; HX44255; HY44255; HZ44255; IA44255; IB44255; IC44255; ID44255; IE44255; IF44255; IG44255; IH44255; II44255; IJ44255; IK44255; IL44255; IM44255; IN44255; IO44255; IP44255; IQ44255; IR44255; IS44255; IT44255; IU44255; IV44255; IW44255; IX44255; IY44255; IZ44255; JA44255; JB44255; JC44255; JD44255; JE44255; JF44255; JG44255; JH44255; JI44255; JJ44255; JK44255; JL44255; JM44255; JN44255; JO44255; JP44255; JQ44255; JR44255; JS44255; JT44255; JU44255; JV44255; JW44255; JX44255; JY44255; JZ44255; KA44255; KB44255; KC44255; KD44255; KE44255; KF44255; KG44255; KH44255; KI44255; KJ44255; KK44255; KL44255; KM44255; KN44255; KO44255; KP44255; KQ44255; KR44255; KS44255; KT44255; KU44255; KV44255; KW44255; KX44255; KY44255; KZ44255; LA44255; LB44255; LC44255; LD44255; LE44255; LF44255; LG44255; LH44255; LI44255; LJ44255; LK44255; LL44255; LM44255; LN44255; LO44255; LP44255; LQ44255; LR44255; LS44255; LT44255; LU44255; LV44255; LW44255; LX44255; LY44255; LZ44255; MA44255; MB44255; MC44255; MD44255; ME44255; MF44255; MG44255; MH44255; MI44255; MJ44255; MK44255; ML44255; MN44255; MO44255; MP44255; MQ44255; MR44255; MS44255; MT44255; MU44255; MV44255; MW44255; MX44255; MY44255; MZ44255; NA44255; NB44255; NC44255; ND44255; NE44255; NF44255; NG44255; NH44255; NI44255; NJ44255; NK44255; NL44255; NM44255; NO44255; NP44255; NQ44255; NR44255; NS44255; NT44255; NU44255; NV44255; NW44255; NX44255; NY44255; NZ44255; OA44255; OB44255; OC44255; OD44255; OE44255; OF44255; OG44255; OH44255; OI44255; OJ44255; OK44255; OL44255; OM44255; ON44255; OO44255; OP44255; OQ44255; OR44255; OS44255; OT44255; OU44255; OV44255; OW44255; OX44255; OY44255; OZ44255; PA44255; PB44255; PC44255; PD44255; PE44255; PF44255; PG44255; PH44255; PI44255; PJ44255; PK44255; PL44255; PM44255; PN44255; PO44255; PP44255; PQ44255; PR44255; PS44255; PT44255; PU44255; PV44255; PW44255; PX44255; PY44255; PZ44255; QA44255; QB44255; QC44255; QD44255; QE44255; QF44255; QG44255; QH44255; QI44255; QJ44255; QK44255; QL44255; QM44255; QN44255; QO44255; QP44255; QQ44255; QR44255; QS44255; QT44255; QU44255; QV44255; QW44255; QX44255; QY44255; QZ44255; RA44255; RB44255; RC44255; RD44255; RE44255; RF44255; RG44255; RH44255; RI44255; RJ44255; RK44255; RL44255; RM44255; RN44255; RO44255; RP44255; RQ44255; RR44255; RS44255; RT44255; RU44255; RV44255; RW44255; RX44255; RY44255; RZ44255; SA44255; SB44255; SC44255; SD44255; SE44255; SF44255; SG44255; SH44255; SI44255; SJ44255; SK44255; SL44255; SM44255; SN44255; SO44255; SP44255; SQ44255; SR44255; SS44255; ST44255; SU44255; SV44255; SW44255; SX44255; SY44255; SZ44255; TA44255; TB44255; TC44255; TD44255; TE44255; TF44255; TG44255; TH44255; TI44255; TJ44255; TK44255; TL44255; TM44255; TN44255; TO44255; TP44255; TQ44255; TR44255; TS44255; TT44255; TU44255; TV44255; TW44255; TX44255; TY44255; TZ44255; UA44255; UB44255; UC44255; UD44255; UE44255; UF44255; UG44255; UH44255; UI44255; UJ44255; UK44255; UL44255; UM44255; UN44255; UO44255; UP44255; UQ44255; UR44255; US44255; UT44255; UV44255; UW44255; UX44255; UY44255; UZ44255; VA44255; VB44255; VC44255; VD44255; VE44255; VF44255; VG44255; VH44255; VI44255; VJ44255; VK44255; VL44255; VM44255; VN44255; VO44255; VP44255; VQ44255; VR44255; VS44255; VT44255; VU44255; VV44255; VW44255; VX44255; VY44255; VZ44255; WA44255; WB44255; WC44255; WD44255; WE44255; WF44255; WG44255; WH44255; WI44255; WJ44255; WK44255; WL44255; WM44255; WN44255; WO44255; WP44255; WQ44255; WR44255; WS44255; WT44255; WU44255; WV44255; WW44255; WX44255; WY44255; WZ44255; XA44255; XB44255; XC44255; XD44255; XE44255; XF44255; XG44255; XH44255; XI44255; XJ44255; XK44255; XL44255; XM44255; XN44255; XO44255; XP44255; XQ44255; XR44255; XS44255; XT44255; XU44255; XV44255; XW44255; XX44255; XY44255; XZ44255; YA44255; YB44255; YC44255; YD44255; YE44255; YF44255; YG44255; YH44255; YI44255; YJ44255; YK44255; YL44255; YM44255; YN44255; YO44255; YP44255; YQ44255; YR44255; YS44255; YT44255; YU44255; YV44255; YW44255; YX44255; YY44255; YZ44255; ZA44255; ZB44255; ZC44255; ZD44255; ZE44255; ZF44255; ZG44255; ZH44255; ZI44255; ZJ44255; ZK44255; ZL44255; ZM44255; ZN44255; ZO44255; ZP44255; ZQ44255; ZR44255; ZS44255; ZT44255; ZU44255; ZV44255; ZW44255; ZX44255; ZY44255; ZZ44255;

A;Title: Primary structure of the mannose receptor contains multiple motifs resembling

A;Reference number: A36563; MUID:90324192; PMID:2373685

A;Accession: A36563

A;Molecule type: mRNA

A;Residues: 1-1456 <TAY>

A;Cross-references: GB:J05550; NID:G188675; PIDN:AAA59868.1; PID:G188676

A;Note: parts of this sequence, including the amino end of the mature protein, were cor

R;Ezekowitz, R.A.B.; Sastri, K.; Bailly, P.; Warner, A.

J. Exp. Med. 172, 1785-1794, 1990

A;Title: Molecular characterization of the human macrophage mannose receptor: demonstra

A;Reference number: A60926; MUID:91079783; PMID:2258707

A;Accession: A60926

A;Status: nucleic acid sequence not shown

A;Molecule type: mRNA

A;Residues: 1-1333,'T',1335-1456 <EZE>

A;Cross-references: GB:X55635

A;Note: translation of the nucleotide sequence is incomplete

A;Note: in the authors' translation additional residues Pro-Glu-Ile are shown after 497

R;Kim, S.J.; Ruiz, N.; Bezouska, K.; Drickamer, K.

Genomics 14, 721-727, 1992

A;Title: Organization of the gene encoding the human macrophage mannose receptor (MRC1)

A;Reference number: A44255; MUID:93052405; PMID:1294118

A;Accession: A44255

A;Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tr

A;Molecule type: DNA

A;Residues: 155-233,'KSAL',238-283;346-428;492-569;631-714,716-719;783-820,'N',822-865;

A;Note: sequence extracted from NCBI backbone (NCBIP:118415, NCBIP:118421, NCBIP:118428

C;Genetics:

A;Gene: GDB:MRC1

A;Cross-references: GDB:133759; OMIM:153618

A;Map position: 10p13-10p13

C;Superfamily: phospholipase A2 receptor; C-type lectin homology; fibronectin type II r

C;Keywords: duplication; lectin; tandem repeat; transmembrane protein

F;1-18/Domain: signal sequence #status predicted <SIG>

F;168-209/Domain: fibronectin type II repeat homology <2F1>

F;223-340/Domain: C-type lectin homology <LCH1>

F;362-486/Domain: C-type lectin homology <LCH2>

F;945-1079/Domain: C-type lectin homology <LCH3>

Query Match

Best Local Similarity 9.2%; Score 183; DB 1; Length 1456;

Matches 53; Conservative 38; Mismatches 66; Indels 54; Gaps 8;

QY 42 YKVIYFHDTSRRLLNFEEAKEACRRDGGQLVSIESEDEQKLIKFIENLLPSDGFWIGLR 101

Db 807 YKDYQYFSKEKETMDNARAFCKRNFGDLVSIQSESEKKFLWKYV-NRNDASAYFIGLL 865
QY 102 RREEKQSNSTACQDLYAWTDGSIQFRNWWYVDEPSCGS--EVCVVMYHQPSAPAGIGGPY 159
Db 866 ISLDKK-----FAWMDGSKVDYVSWATGEPNFANEDENCVTMY-----SNSGF---- 908
QY 160 MFQWNDDRCNMKNFNICKYSDEK----PAVPSREAEGETELTPVLPEETQE----- 208
Db 909 ---WNDINCGYPNAFICQRHNSSINATVMP-----TMSVPVSGCKEGWNFYSN 954
QY 209 -----EDAKKTFKESREAAALNLAYILI 230
Db 955 KCFKIFGFMEEERKNWQEARAKACIGFGGNLV 985
RESULT 2
mannose receptor precursor, macrophage - mouse
C:Species: Mus musculus (house mouse)
C:Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A48925; S21320; PC2245
R:Harris, N.; Super, M.; Rits, M.; Chang, G.; Ezekowitz, R.A.
Blood 80, 2363-2373, 1992
A:Title: Characterization of the murine macrophage mannose receptor: demonstration that
A:Reference number: A48925; MUID:93043353; PMID:1421407
A:Accession: A48925
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-1455 <HAR>
A:Experimental source: peritoneal macrophage
A:Note: sequence extracted from NCBI backbone (NCBIP:118733)
R:Harris, N.; Super, M.; Rits, M.; Chang, G.; Ezekowitz, R.B.
submitted to the EMBL Data Library, April 1992
A:Description: Characterization of the murine macrophage mannose receptor: Demonstration
on.
A:Reference number: S21320
A:Accession: S21320
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-302,'W',303-1117,'E',1119-1455 <HA2>
A:Cross-references: EMBL:Z11974; NID:G52997; PIDN:CAA78028.1; PID:G52998
R:Harris, N.; Peters, L.L.; Eicher, E.M.; Rits, M.; Raspberry, D.; Eichbaum, Q.G.; Super
Biochem. Biophys. Res. Commun. 198, 682-692, 1994
A:Title: The exon-intron structure and chromosomal localization of the mouse macrophage
A:Reference number: PC2245; MUID:94128116; PMID:8297379
A:Accession: PC2245
A:Molecule type: mRNA
A:Residues: 35-105 <HA3>
C:Genetics:
A:Gene: Mrcl
A:Map position: 2
C:Superfamily: phospholipase A2 receptor; C-type lectin homology; fibronectin type II re
C:Keywords: membrane protein; receptor
F:168-209/Domain: fibronectin type II repeat homology <2F9>
F:361-485/Domain: C-type lectin homology <LCH1>
F:943-1077/Domain: C-type lectin homology <LCH2>
Query Match 8.9%; Score 178.5; DB 1; Length 1455;
Best Local Similarity 25.9%; Pred. No. 2.7e-06;
Matches 53; Conservative 33; Mismatches 66; Indels 53; Gaps 8;
QY 42 YKVIYFHDTSRRLNFEAEKACRRDGGQLVSIQSESEKFLWKYI-NKNGGQSPYFIGML 101
Db 806 YKDYQYFSKEKETMDNARRFCNKGFDLATIKSESEKFLWKYI-NKNGGQSPYFIGML 864
QY 102 RREEKQSNSTACQDLYAWTDGSIQFRNWWYVDEPSCGS--EVCVVMYHQPSAPAGIGGPY 159
Db 865 ISMDKK-----FIWMDGSKVDFVAWATGEPNFANDDENCVTMY-----TNSGF---- 907
QY 160 MFQWNDDRCNMKNFNICK----YSDEKPAVPSREAEGETELTPVLPEETQE----- 208
Db 908 ---WNDINCGYPNFIQCRHNSSINATAMP-----TTPTPGCKEGWHLKYNK 953

QY 209 -----EDAKKTFKESREAAALNL 225
Db 954 CFKIFGFANEKKSWQDARQACKGL 978
RESULT 3
S52781
neurocan - mouse
C:Species: Mus musculus (house mouse)
C:Date: 19-May-1995 #sequence revision 21-Jul-1995 #text_change 04-Feb-2000
C:Accession: S52781
R:Rauch, U.; Forsberg, N.; Kulbe, G.; Arnold-Ammer, I.; Faessler, R.
submitted to the EMBL Data Library, February 1995
A:Description: Amino acid sequence of mouse neurocan and brevican and their different e
A:Reference number: S52781
A:Accession: S52781
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1268 <RAU>
A:Cross-references: EMBL:X84727; NID:G758629; PIDN:CAA59216.1; PID:G758630
C:Superfamily: aggrecan; C-type lectin homology; complement factor H repeat homology; E
F:176-253/Domain: link protein repeat homology <LNK1>
F:274-355/Domain: link protein repeat homology <LNK2>
F:964-995/Domain: EGF homology <EGF>
F:1040-1160/Domain: C-type lectin homology <LCH>
F:1167-1223/Domain: complement factor H repeat homology <FHD>
Query Match 8.8%; Score 177; DB 2; Length 1268;
Best Local Similarity 31.5%; Pred. No. 3.1e-06;
Matches 45; Conservative 17; Mismatches 49; Indels 32; Gaps 6;
QY 38 QRPCYKVIYFHDTSRRLNFEAEKACRRDGGQLVSIQSESEKFLWKYI-NKNGGQSPYFIGML 97
Db 1048 QGHCYR--YF---AHRRAWEDAERDCRRRAGHLTSVHSPBEHKFINSF-----GHENSW 1096
QY 98 IGLRRREKQSNSTACQDLYAWTDGSIQFRNWWYVDEPS---CGSEVCVVMYHQPSAPAG 154
Db 1097 IGLNDRTVTRD-----FQMTDNTGLQYENWREKQPDNFFAGGEDCVVMVAHESG--- 1145
QY 155 IGGPYMFQWNDDRCNMKNFNICK 177
Db 1146 -----RWNDVPCNYNLPYVCK 1161
RESULT 4
T14274
versican precursor, splice form V2 - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 20-Sep-1999 #sequence revision 20-Sep-1999 #text_change 05-May-2000
C:Accession: T14274
R:Schmalfeldt, M.; Dours-Zimmermann, M.T.; Winterhalter, K.H.; Zimmermann, D.R.
J. Biol. Chem. 273, 15758-15764, 1998
A:Title: Versican V2 is a major extracellular matrix component of the mature bovine bra
A:Reference number: 217954; MUID:98288320; PMID:9624174
A:Accession: T14274
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1643 <SCH>
A:Cross-references: EMBL:AF060458; NID:G3253303; PID:G3253304; PIDN:AAC24360.1
A:Experimental source: brain
C:Keywords: glycoprotein
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-1643/Product: versican, splice form V2 #status predicted <MAT>
F:57,331,352,817,965,1017,1333,1616,1626/Binding site: carbohydrate (asn) (covalent) #8
Query Match 8.7%; Score 174.5; DB 2; Length 1643;
Best Local Similarity 25.5%; Pred. No. 6.8e-06;
Matches 59; Conservative 26; Mismatches 71; Indels 75; Gaps 11;
QY 38 QRPCYKVIYFHDTSRRLNFEAEKACRRDGGQLVSIQSESEKFLWKYI-NKNGGQSPYFIGML 96
Db 1424 QGQCYK--YF---AHRRTWDAARECRLOGAHLTSLSHEEQMFVNRV-----GHYDQ 1471

RESULT 8
A60979
versican precursor - human
N:Alternate names: chondroitin sulfate proteoglycan 2; chondroitin sulfate proteoglycan
N:Contains: glial hyaluronate-binding protein
C:Species: Homo sapiens (man)
C:Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text change 19-Jan-2001
C:Accession: S06014; S43921; A60979; A30358; A29348; A45131; I54179
R:Zimmermann, D.R.; Ruoslahti, E.
EMBO J. 8, 2975-2981, 1989
A:Title: Multiple domains of the large fibroblast proteoglycan, versican.
A:Reference number: S06014; MUID:90059882; PMID:2583089
A:Accession: S06014
A:Molecule type: mRNA
A:Residues: 1-2409 <ZIM>
A:Cross-references: GB:X15998; NID:G37662; PIDN:CAA34128.1; PID:G37663
R:Yao, L.Y.; Moody, C.; Schoenherr, E.; Wight, T.N.; Sandell, L.J.
Matrix Biol. 14, 213-225, 1994
A:Title: Identification of the proteoglycan versican in aorta and smooth muscle cells by
A:Reference number: S43921; MUID:95005762; PMID:7921538
A:Accession: S43921
A:Molecule type: mRNA
A:Residues: 208-440;1094-1385;1910-2246 <YAO>
R:Bignami, A.; Lane, W.S.; Andrews, D.; Dahl, D.
Brain Res. Bull. 22, 67-70, 1989
A:Title: Structural similarity of hyaluronate binding proteins in brain and cartilage.
A:Reference number: A60979; MUID:89229983; PMID:2469524
A:Accession: A60979
A:Molecule type: protein
A:Residues: 171-210;289-303 <BIG>
R:Perides, G.; Lane, W.S.; Andrews, D.; Dahl, D.; Bignami, A.
J. Biol. Chem. 264, 5981-5987, 1989
A:Title: Isolation and partial characterization of a glial hyaluronate-binding protein.
A:Reference number: A30358; MUID:89174663; PMID:2466833
A:Accession: A30358
A:Molecule type: protein
A:Residues: 24-50;80-87,'D',89-119;128-155;167-218;229-259,'IR',261-268;277-283,'G',285-
R:Krusius, T.; Gehlsen, K.R.; Ruoslahti, E.
J. Biol. Chem. 262, 13120-13125, 1987
A:Title: A fibroblast chondroitin sulfate proteoglycan core protein contains lectin-like
A:Reference number: A29348; MUID:88007514; PMID:2820964
A:Accession: A29348
A:Molecule type: mRNA
A:Residues: 1725,'V',1727-2409 <KRU>
A:Cross-references: GB:J02814
R:Perides, G.; Rahemtulla, F.; Lane, W.S.; Asher, R.A.; Bignami, A.
J. Biol. Chem. 267, 23883-23887, 1992
A:Title: Isolation of a large aggregating proteoglycan from human brain.
A:Reference number: A45131; MUID:93054750; PMID:1429726
A:Contents: brain
A:Accession: A45131
A:Molecule type: protein
A:Residues: 21-22,'X',24-37 <PE2>
A:Experimental source: brain
A:Note: sequence extracted from NCBI backbone (NCBIP:118884)
R:Iozzo, R.V.; Naso, M.F.; Cannizzaro, L.A.; Wasmuth, J.J.; McPherson, J.D.
Genomics 14, 845-851, 1992
A:Title: Mapping of the versican proteoglycan gene (CSPG2) to the long arm of human chro
A:Reference number: I54179; MUID:93122792; PMID:1478664
A:Accession: I54179
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 251-347 <RES>
A:Cross-references: GB:S52488; NID:G263313; PIDN:AAB24878.1; PID:G263314
C:Genetics:
A:Gene: GDB:CSPG2
A:Cross-references: GDB:127873; OMIM:118661
A:Map position: 5q12-5q14
C:Superfamily: versican; C-type lectin homology; complement factor H repeat homology; EG
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-2409/Product: proteoglycan 24K core protein #status predicted <MAT>
F:167-244/Domain: link protein repeat homology <LNK1>

F:265-346/Domain: link protein repeat homology <LNK2>
F:559-1654/Domain: chondroitin sulfate attachment #status predicted <GAG>
F:2106-2137/Domain: EGF homology <EG1>
F:2144-2175/Domain: EGF homology <EG2>
F:2182-2302/Domain: C-type lectin homology <LCH>
F:2309-2365/Domain: complement factor H repeat homology <FHD>

Query Match 8.7%; Score 174; DB 1; Length 2409;
Best Local Similarity 28.5%; Pred. No. 1.2e-05;
Matches 47; Conservative 23; Mismatches 55; Indels 40; Gaps 8;

QY 38 QRPCYKVIYFHDTSRRLLNFEEAKEACRRDGGQLVSI ESEDEQKLI EKFIENLLPSD GDF- 96
| | | | | : : : : : | | | | | : : : : :
Db 2190 QGQCYK--YF---AHRRTWDAARECRLQGAHLT SILSHEEQVFVNRV-----GHDYQ 2237
| | | | | : : : : : | | | | | : : : : :

QY 97 WIGLRRREEKQSNSTACQDLYAWTDGSI SQFRN WYVDEP----SCGSEVCVVMYHQPSAP 152
| | | | | : : : : : | | | | | : : : : :
Db 2238 WIGL-----NDKMFEDFRWTDGSTLQYENWRPNQPDSPFS SAGEDCVV IWHENG-- 2287
| | | | | : : : : : | | | | | : : : : :

QY 153 AGIGGPFYMFQWDDRCNMKNFICKYS----DEKPAVPSREAEGE 193
| | | | | : : : : : | | | | | : : : : :
Db 2288 -----QWNDVPCNYHLTYTCKKGTVACGQPPVVENAKTFGK 2323
| | | | | : : : : : | | | | | : : : : :

RESULT 9
A47171
chondroitin sulfate proteoglycan PG-M core protein - chicken
C:Species: Gallus gallus (chicken)
C:Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 21-Jul-2000
C:Accession: A47171
R:Shinomura, T.; Nishida, Y.; Ito, K.; Kimata, K.
J. Biol. Chem. 268, 14461-14469, 1993
A:Title: cDNA cloning of PG-M, a large chondroitin sulfate proteoglycan expressed durin

A:Reference number: A47171; MUID:93300846; PMID:8314802
A:Accession: A47171
A:Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-3562 <SHI>
A:Cross-references: GB:D13542; NID:G391643; PIDN:BAA02742.1; PID:G391644
A:Experimental source: stage 22-23 developing limb buds
A:Note: sequence extracted from NCBI backbone (NCBIN:134456, NCBIP:134457)
C:Superfamily: chicken chondroitin sulfate proteoglycan PG-M core protein; C-type lectin
F:166-243/Domain: link protein repeat homology <LNK1>
F:264-345/Domain: link protein repeat homology <LNK2>
F:3258-3289/Domain: EGF homology <EGF1>
F:3296-3327/Domain: EGF homology <EGF>
F:3334-3454/Domain: C-type lectin homology <LCH>
F:3461-3517/Domain: complement factor H repeat homology <FHD>

Query Match 8.6%; Score 171; DB 2; Length 3562;
Best Local Similarity 28.5%; Pred. No. 3.5e-05;
Matches 47; Conservative 23; Mismatches 55; Indels 40; Gaps 8;

QY 38 QRPCYKVIYFHDTSRRLLNFEEAKEACRRDGGQLVSI ESEDEQKLI EKFIENLLPSD GDF- 96
| | | | | : : : : : | | | | | : : : : :
Db 3342 QGQCYK--YF---AHRRTWDTARECRLQGAHLT SILSHEEQVFVNR1-----GHDYQ 3389
| | | | | : : : : : | | | | | : : : : :

QY 97 WIGLRRREEKQSNSTACQDLYAWTDGSI SQFRN WYVDEP----SCGSEVCVVMYHQPSAP 152
| | | | | : : : : : | | | | | : : : : :
Db 3390 WIGL-----NDKMFEDFRWTDGSTLQYENWRPNQPDSPFS SAGEDCVV IWHENG-- 3439
| | | | | : : : : : | | | | | : : : : :

QY 153 AGIGGPFYMFQWDDRCNMKNFICKYS----DEKPAVPSREAEGE 193
| | | | | : : : : : | | | | | : : : : :
Db 3440 -----QWNDVPCNYHLTYTCKKGTVACGQPPVVENAKTFGK 3475
| | | | | : : : : : | | | | | : : : : :

RESULT 10
T42710
mannose receptor, macrophage - mouse
N:Alternate names: lambda lectin; phospholipase A2 receptor
C:Species: Mus musculus (house mouse)
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jun-2000

C;Accession: T42710
R;Wu, K.; Yuan, J.; Lasky, L.A.
J. Biol. Chem. 271, 21323-21330, 1996
A;Title: Characterization of a novel member of the macrophage mannose receptor type C le
A;Reference number: 222235; MUID:96355501; PMID:8702911
A;Accession: T42710
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1479 <WUK>
A;Cross-references: EMBL:U56734; NID:G1336073; PID:G1336074; PIDN:AAC52729.1
C;Superfamily: phospholipase A2 receptor; C-type lectin homology; fibronectin type II re
C;Keywords: membrane protein; receptor
F;186-227/Domain: fibronectin type II repeat homology <2FR>

Query Match 8.5%; Score 170.5; DB 2; Length 1479;
Best Local Similarity 31.4%; Pred. No. 1.3e-05;
Matches 48; Conservative 22; Mismatches 48; Indels 35; Gaps 7;

QY 28 SQPVCRCGTQPCYKVIYFHDTSRRLNFEAEKACRRDGGQLVSIIESEDEQKLIIEKFI 87
Db 384 SWQPF-----QGHCYRL-----QAERKRWQESKRACLRGGDLISHSMAELEFITKQIK 433

QY 88 NLLPSDGFWIGLRRREKQSNSTACQDLYAWTDGSIQFRNWWYVDEPS---CGSEVCVV 144
Db 434 QEVE---ELWIGL-----NDLKLQMNFEWSDGLSVFTWHPPFPNFRDLSLEDCVT 482

QY 145 MYHQPAPAGIGGYPMFQWDDRCNMKNFICK 177
Db 483 IW----GPEG-----RWNDSPCNQSLPSICK 504

RESULT 11
A39808
proteoglycan core protein, cartilage - bovine (fragments)
N;Alternate names: aggrecan; aggregating cartilage proteoglycan
C;Species: Bos primigenius taurus (cattle)
C;Date: 20-Mar-1992 #sequence_revision 23-Mar-1995 #text_change 13-Aug-1999
C;Accession: A34234; A27752; A39808; A27751; E29164; B27751; D27751; E27751; F27
R;Antonsson, P.; Heinegard, D.; Oldberg, A.
J. Biol. Chem. 264, 16170-16173, 1989
A;Title: The keratan sulfate-enriched region of bovine cartilage proteoglycan consists o
A;Reference number: A34234; MUID:89380219; PMID:2528543
A;Accession: A34234
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 128-621 <ANT>
A;Cross-references: GB:J05028
R;Oldberg, A.; Antonsson, P.; Heinegard, D.
Biochem. J. 243, 255-259, 1987
A;Title: The partial amino acid sequence of bovine cartilage proteoglycan, deduced from a
A;Reference number: A27752; MUID:87270630; PMID:3111460
A;Accession: A27752
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 622-1340 <OLD>
R;Sandy, J.D.; Boynton, R.E.; Flannery, C.R.
J. Biol. Chem. 266, 8198-8205, 1991
A;Title: Analysis of the catabolism of aggrecan in cartilage explants by quantitation of
A;Reference number: A39808; MUID:91217051; PMID:2022637
A;Accession: A39808
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-28;59-82;131-137,'QSET',142-149;196-207;226-249;1137-1143;1252-1267;1274-1
R;Perin, J.P.; Bonnet, F.; Jolles, P.
FEBS Lett. 206, 73-77, 1986
A;Title: Structural relationship between link proteins and proteoglycan monomers.
A;Reference number: A27751; MUID:87005253; PMID:3530809
A;Accession: A27751
A;Molecule type: protein
A;Residues: 29-58;74-130;174-175,'A',177-204;208-225 <PER>
R;Perin, J.P.; Bonnet, F.; Jolles, J.; Jolles, P.
FEBS Lett. 176, 37-42, 1984
A;Title: Sequence data concerning the protein core of the cartilage proteoglycan monomer

A;Reference number: A91327; MUID:85027710; PMID:6489519
A;Accession: E29164
A;Molecule type: protein
A;Residues: 1230-1249 <PE2>
C;Superfamily: aggrecan; C-type lectin homology; complement factor H repeat homology;
C;Keywords: glycoprotein
F;8-28/Domain: link protein repeat homology (fragment) <LNK1>
F;29-58/Domain: link protein repeat homology (fragment) <LNK2>
F;80-146/Domain: link protein repeat homology (fragments) <LNK3>
F;167-248/Domain: link protein repeat homology <LNK4>
F;1130-1250/Domain: C-type lectin homology <LCH>
F;1257-1313/Domain: complement factor H repeat homology <FHD>

Query Match 7.9%; Score 158.5; DB 2; Length 1340;
Best Local Similarity 26.8%; Pred. No. 0.00011;
Matches 48; Conservative 29; Mismatches 59; Indels 43; Gaps 10;

QY 30 QPVCRCG---TORPCYKVIYFHDTSRRLNFEAEKACRRDGGQLVSIIESEDEQKLIIEKFI 86
Db 1127 QKLCBEGWTKFQHCYR--HFPD---RATWVDAESQCRKQKQSHLSSIVTPEEQ----EFV 1177

QY 87 ENLLPSDGF-WIGLRRREKQSNSTACQDLYAWTDGSIQFRNWWYVDEP----SCGSEV 141
Db 1178 NN---NAQDYQWIGL-----NDKTIEGDFRWSHGSLQFENWRPNQDNFFATGEDC 1226

QY 142 CVVMYHQPAPAGIGGYPMFQWDDRCNMKNFICKYS----DEKPAVPSREAEGETE 196
Db 1227 VVMIWHEKG-----EWNDVPCNYQLPFTCKKGTVACGPPVVEHARIFGQKKD 1274

RESULT 12
T42630

aggrecan - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000
C;Accession: T42630
R;Hering, T.M.; Kollar, J.; Huynh, T.D.
submitted to the EMBL Data Library, September 1996
A;Description: Complete coding sequence of bovine aggrecan: comparative structural ana
A;Reference number: Z22182
A;Accession: T42630
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-2327 <HER>
A;Cross-references: EMBL:U76615; NID:G1730259; PID:G1730260; PIDN:AAB38524.1
A;Experimental source: articular chondrocytes
C;Superfamily: aggrecan; C-type lectin homology; complement factor H repeat homology;
C;Keywords: cartilage; chondroitin sulfate proteoglycan; extracellular matrix; glycopr

Query Match 7.9%; Score 158.5; DB 2; Length 2327;
Best Local Similarity 26.8%; Pred. No. 0.00022;
Matches 48; Conservative 29; Mismatches 59; Indels 43; Gaps 10;

QY 30 QPVCRCG---TORPCYKVIYFHDTSRRLNFEAEKACRRDGGQLVSIIESEDEQKLIIEKFI 86
Db 2114 QKLCBEGWTKFQHCYR--HFPD---RATWVDAESQCRKQKQSHLSSIVTPEEQ----EFV 2164

QY 87 ENLLPSDGF-WIGLRRREKQSNSTACQDLYAWTDGSIQFRNWWYVDEP----SCGSEV 141
Db 2165 NN---NAQDYQWIGL-----NDKTIEGDFRWSHGSLQFENWRPNQDNFFATGEDC 2213

QY 142 CVVMYHQPAPAGIGGYPMFQWDDRCNMKNFICKYS----DEKPAVPSREAEGETE 196
Db 2214 VVMIWHEKG-----EWNDVPCNYQLPFTCKKGTVACGPPVVEHARIFGQKKD 2261

RESULT 13
A39086

aggrecan precursor, cartilage long splice form [validated] - human
N;Alternate names: chondroitin sulfate proteoglycan 1; large aggregating proteoglycan;
N;Contains: aggrecan cartilage short splice form
C;Species: Homo sapiens (man)
C;Date: 10-Sep-1999 #sequence_revision 01-Dec-2000 #text_change 08-Dec-2000

F;26-150/Domain: C-type lectin homology <LCH>
F;26-39,56-150,125-142/Disulfide bonds: #status experimental
F;157/Disulfide bonds: interchain (to 160) #status experimental
F;160/Disulfide bonds: interchain (to 157) #status experimental

Query Match 7.8%; Score 156; DB 1; Length 162;
Best Local Similarity 24.9%; Pred. No. 1.3e-05;
Matches 46; Conservative 30; Mismatches 67; Indels 42; Gaps 7;

Qy 1 MRPGTALQAVLLVGLRAATGRLLSGQVPCRGGTQPCYKVIYFHDTSRRLNPFEEAK 60
Db 1 MQRSEIVQAVTLVAVFA-----ITGECTCPGNLDWQYDGHCIWASTYQVRWDAQ 53

Qy 61 EACR--RDGGQLVSIIESEDEQKLIKFIENLLPSDGDGFWIGLRRREKQSNSTACQDLYA 118
Db 54 LACQTVHPGAYLATIQSLENAFISQTVSN-----NRLWIGL-----NDIDLEGHYV 100

Qy 119 WTDGSIQFRNWWYVDEPS-----CGSEVVCVMYHQPSAPAGIGGPPYMFQWNDRCNMKN 172
Db 101 WSGEATDFTYWSSNNPNWENQDCG----VVNYDTVTG-----QWDDDDCNKNR 146

Qy 173 NFICK 177
Db 147 NFLCK 151

RESULT 15
B42755
E-selectin precursor - mouse
N;Alternate names: endothelial leukocyte adhesion molecule 1 (ELAM-1)
C;Species: Mus musculus (house mouse)
C;Date: 04-Mar-1993 #sequence_revision 31-Dec-2000 #text_change 02-Aug-2002
C;Accession: S23174; B42755
R;Becker-Andre, M.; van Huijsduijn, R.H.; Losberger, C.; Whelan, J.; Delamarter, J.F.
Eur. J. Biochem. 206, 401-411, 1992
A;Title: Murine endothelial leukocyte-adhesion molecule 1 is a close structural and functional homologue of the mouse endothelial selectins. Expression of both E- and P-selectin in murine endothelial cells.
A;Reference number: S23174; MUID:92283265; PMID:1375914
A;Accession: S23174
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-612 <BEC>
A;Cross-references: GB:M80778; NID:g193014; PIDN:AAA37547.1; PID:g193015
R;Weller, A.; Isenmann, S.; Vestweber, D.
J. Biol. Chem. 267, 15176-15183, 1992
A;Title: Cloning of the mouse endothelial selectins. Expression of both E- and P-selectin in murine endothelial cells.
A;Reference number: A42755; MUID:92340571; PMID:1378846
A;Accession: B42755
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 'MKATAGV', 1-389, 391-612 <WEL>
A;Cross-references: GB:M87862; NID:g193107
A;Experimental source: endothelial cells
A;Note: sequence extracted from NCBI backbone (NCBIP:109470)
A;Note: the sequence in GenBank entry MUSESELEC, release 117.0, (PIDN:AAA37577.1; PID:g193107) is uncertain whether the initiator is Met-1 or the AUG codon preceding that
C;Superfamily: P-selectin; C-type lectin homology; complement factor H repeat homology;
C;Keywords: glycoprotein; transmembrane protein
F;1-21/Domain: signal sequence #status predicted <SIG>
F;12-138/Domain: C-type lectin homology <LCH>
F;22-612/Product: P-selectin #status predicted <MAT>
F;143-174/Domain: EGF homology <EGF>
F;180-238/Domain: complement factor H repeat homology <FH1>
F;243-300/Domain: complement factor H repeat homology <FH2>
F;305-363/Domain: complement factor H repeat homology <FH3>
F;368-426/Domain: complement factor H repeat homology <FH4>
F;431-489/Domain: complement factor H repeat homology <FH5>
F;494-548/Domain: complement factor H repeat homology <FH6>
F;25,391,528/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 7.7%; Score 154.5; DB 2; Length 612;
Best Local Similarity 27.9%; Pred. No. 9e-05;
Matches 38; Conservative 28; Mismatches 45; Indels 25; Gaps 5;

Qy 46 YFHDTSRRLNFEAKEACRRDGGQLVSIIESEDEQKLIKFIENLLPSDGDGFWIGLRRREE 105
Db 23 YYNASSELMTYDEASAYCQRDYTHLVAIQNKEB---INVLNSNLKHSPSYWIGIRK--- 76

Qy 106 KQSNSTACQDLYAWTDGS---ISQFRNWWYVDEPS---CGSEVVCVMYHQPSAPAGIGGPPY 160
Db 77 -----VNNVWIIWVGTKPLTEEAQNWAPGEPNNKQNRNEDCVBIYIORTKOSGM----- 124

Qy 161 FQWNDRCNMKNFIC' 176
Db 125 --WDERCNKKLALC 138

Search completed: December 22, 2003, 16:14:40
Job time : 20.1827 secs


```
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-909-320-137

Query Match      99.3%; Score 1986; DB 10; Length 382;
Best Local Similarity 97.9%; Pred. No. 2.7e-178;
Matches 374; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 MRPGTALQAVLLAVLLVGLRAATGRLLS-----GQVCRGGGTQPCYKVIYFHDTSR 52
Db 1 MRPGTALQAVLLAVLLVGLRAATGRLLSASDLDLRGQGPVCRGGGTQPCYKVIYFHDTSR 60

QY 53 RLNFEEAKEACRRDGGQLVSIESTEDEQKLIKFIENLLPSDGFWIGLRRRREEKQSNSTA 112
Db 61 RLNFEEAKEACRRDGGQLVSIESTEDEQKLIKFIENLLPSDGFWIGLRRRREEKQSNSTA 120

QY 113 CQDLyawtdgsisqfrnwyvdepscgsevcvmyhqsapagiggympfQWNNDRCNMKN 172
Db 121 CQDLyawtdgsisqfrnwyvdepscgsevcvmyhqsapagiggympfQWNNDRCNMKN 180

QY 173 NFICKYSDEKPAVPSREAEGETELTTPVLPEETQEEEDAKKTFKESREAAALNLAYILIPS 232
Db 181 NFICKYSDEKPAVPSREAEGETELTTPVLPEETQEEEDAKKTFKESREAAALNLAYILIPS 240

QY 233 IPLLlLLVVTTVVCWVWICRKRKREQDPDPSTKKQHTIWPSPHQGNSPDLEVNVIRKQSE 292
Db 241 IPLLlLLVVTTVVCWVWICRKRKREQDPDPSTKKQHTIWPSPHQGNSPDLEVNVIRKQSE 300

QY 293 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFTNDIYE 352
Db 301 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFTNDIYE 360

QY 353 FSPDQMGRSKESGWENEIYGY 374
Db 361 FSPDQMGRSKESGWENEIYGY 382
```

RESULT 6

```
US-09-909-088B-137
; Sequence 137, Application US/09909088B
; Patent No. US20020146709A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
```

```
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/909,088B
; CURRENT FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-909-088B-137
```

```
Query Match      99.3%; Score 1986; DB 10; Length 382;
Best Local Similarity 97.9%; Pred. No. 2.7e-178;
Matches 374; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 MRPGTALQAVLLAVLLVGLRAATGRLLS-----GQVCRGGGTQPCYKVIYFHDTSR 52
Db 1 MRPGTALQAVLLAVLLVGLRAATGRLLSASDLDLRGQGPVCRGGGTQPCYKVIYFHDTSR 60

QY 53 RLNFEEAKEACRRDGGQLVSIESTEDEQKLIKFIENLLPSDGFWIGLRRRREEKQSNSTA 112
Db 61 RLNFEEAKEACRRDGGQLVSIESTEDEQKLIKFIENLLPSDGFWIGLRRRREEKQSNSTA 120

QY 113 CQDLyawtdgsisqfrnwyvdepscgsevcvmyhqsapagiggympfQWNNDRCNMKN 172
Db 121 CQDLyawtdgsisqfrnwyvdepscgsevcvmyhqsapagiggympfQWNNDRCNMKN 180

QY 173 NFICKYSDEKPAVPSREAEGETELTTPVLPEETQEEEDAKKTFKESREAAALNLAYILIPS 232
Db 181 NFICKYSDEKPAVPSREAEGETELTTPVLPEETQEEEDAKKTFKESREAAALNLAYILIPS 240

QY 233 IPLLlLLVVTTVVCWVWICRKRKREQDPDPSTKKQHTIWPSPHQGNSPDLEVNVIRKQSE 292
Db 241 IPLLlLLVVTTVVCWVWICRKRKREQDPDPSTKKQHTIWPSPHQGNSPDLEVNVIRKQSE 300

QY 293 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFTNDIYE 352
Db 301 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFTNDIYE 360

QY 353 FSPDQMGRSKESGWENEIYGY 374
```


Db 361 FSPDQMGRSKESGWVENEIYG 382
RESULT 7
US-09-905-291A-137
; Sequence 137, Application US/09905291A
; Patent No. US20020160374A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/905,291A
; CURRENT FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05

; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-905-291A-137
Query Match 99.3%; Score 1986; DB 10; Length 382;
Best Local Similarity 97.9%; Pred. No. 2.7e-178;
Matches 374; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
Qy 1 MRPGTALQAVLLAVLLVGLRAATGRLLS-----GQVCRGGTQPCYKVIYFHDTSR 52
Db 1 MRPGTALQAVLLAVLLVGLRAATGRLLSASDLRLRGQGPVCRGGTQPCYKVIYFHDTSR 60
Qy 53 RLNFEEAKEACRRDGGQLVSISEDEQKLIKFIENLLPSDGFWMGLRRREEKQSNSTA 112
Db 61 RLNFEEAKEACRRDGGQLVSISEDEQKLIKFIENLLPSDGFWMGLRRREEKQSNSTA 120
Qy 113 CODLYAWTDGSIQFRNWWYVDEPPSCGSEVCVVMYHQPSAPAGIGGPPYMFQWDDRCNMKN 172
Db 121 CODLYAWTDGSIQFRNWWYVDEPPSCGSEVCVVMYHQPSAPAGIGGPPYMFQWDDRCNMKN 180
Qy 173 NFICKYSDEKPAVPSREAEGETELTTPVLPEETQEEEDAKKTFKESREAAALNLAYILIPS 232
Db 181 NFICKYSDEKPAVPSREAEGETELTTPVLPEETQEEEDAKKTFKESREAAALNLAYILIPS 240
Qy 233 IPLLILLVTTTVCWWVICRKRKREQDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSE 292
Db 241 IPLLILLVTTTVCWWVICRKRKREQDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSE 300
Qy 293 ADLAETRPDLKNIISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVEGFTNDIYE 352
Db 301 ADLAETRPDLKNIISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVEGFTNDIYE 360
Qy 353 FSPDQMGRSKESGWVENEIYG 374
Db 361 FSPDQMGRSKESGWVENEIYG 382
RESULT 8
US-09-902-853-137
; Sequence 137, Application US/09902853
; Publication No. US20020192659A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14

; CURRENT APPLICATION NUMBER: US/09/902,853
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US/09/665,350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-902-853-137

Query Match 99.3%; Score 1986; DB 10; Length 382;
Best Local Similarity 97.9%; Pred. No. 2.7e-178;
Matches 374; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 MRPGTALQAVLLAVLLVGLRAATGRLLS-----GQPVCRGGTQRPCVKYVFHDTSR 52
Db 1 MRPGTALQAVLLAVLLVGLRAATGRLLSASDLRGGQPVCRGGTQRPCVKYVFHDTSR 60
QY 53 RLNFEEAKEACRRDGGQLVSIESEDEQKLIKFIEFNLPSDGFWIGLRREEKQSNSTA 112
Db 61 RLNFEEAKEACRRDGGQLVSIESEDEQKLIKFIEFNLPSDGFWIGLRREEKQSNSTA 120
QY 113 QODLYAWTDGSIQFRNWIYVDEPSCGSEVCVVMYHQPSAPAGIGGPFQWDDRCNMKN 172
Db 121 QODLYAWTDGSIQFRNWIYVDEPSCGSEVCVVMYHQPSAPAGIGGPFQWDDRCNMKN 180
QY 173 NFICKYSDEKPAVPSREAEGETELTTPVLPETQEEDAKKTFKESREAAINLAYILIPS 232
Db 181 NFICKYSDEKPAVPSREAEGETELTTPVLPETQEEDAKKTFKESREAAINLAYILIPS 240
QY 233 IPLLLLLVTTTVCVWVICRKRKEQPDPTKQHTIWPSPHQNSPDLEVNVIRKQSE 292
Db 241 IPLLLLLVTTTVCVWVICRKRKEQPDPTKQHTIWPSPHQNSPDLEVNVIRKQSE 300
QY 293 ADLAETRPDLKNISFRVCSGEATPDDMSCDYNMVA NPSESGFVTLVSVESGFVTNDIYE 352
Db 301 ADLAETRPDLKNISFRVCSGEATPDDMSCDYNMVA NPSESGFVTLVSVESGFVTNDIYE 360
QY 353 FSPDQMGRSKESGWVENEIYGY 374
|

Db 361 FSPDQMGRSKESGWVENEIYGY 382
RESULT 9
US-09-907-824-137
; Sequence 137, Application US/09907824
; Publication No. US20020197671A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/907,824
; CURRENT FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: 09/665,350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219

;
;
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-907-824-137

Query Match 99.3%; Score 1986; DB 10; Length 382;
Best Local Similarity 97.9%; Pred. No. 2.7e-178;
Matches 374; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 MRPGTALQAVLLAVLLVGLRAATGRLLS-----GQVCRGGTQRPCYKVIYFHDTSR 52
Db 1 MRPGTALQAVLLAVLLVGLRAATGRLLSASDLRLGGQPVCRGGTQRPCYKVIYFHDTSR 60

QY 53 RLNFEEAKEACRRDGGQLVSI ESEDEQKLI EKFIENLLPSD GDFWIGLRREEKQSNSTA 112
Db 61 RLNFEEAKEACRRDGGQLVSI ESEDEQKLI EKFIENLLPSD GDFWIGLRREEKQSNSTA 120

QY 113 CQDLYAWTDGSI SQFRNMYVDEPSCGSEVCVVMYHQPSAPAGIGGPFYQWDDRCNMKN 172
Db 121 CQDLYAWTDGSI SQFRNMYVDEPSCGSEVCVVMYHQPSAPAGIGGPFYQWDDRCNMKN 180

QY 173 NFICKYSDEKPAVPSREAEGETELTTPVLPEETOEDAKKTFKESREAAALNLAYILIPS 232
Db 181 NFICKYSDEKPAVPSREAEGETELTTPVLPEETOEDAKKTFKESREAAALNLAYILIPS 240

QY 233 IPLL LLLVVTTVVCWVICRKRKREQPDSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSE 292
Db 241 IPLL LLLVVTTVVCWVICRKRKREQPDSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSE 300

QY 293 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYE 352
Db 301 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYE 360

QY 353 FSPDQMGRSKESGWVENEIYGY 374
Db 361 FSPDQMGRSKESGWVENEIYGY 382

RESULT 10

US-09-907-841-137
; Sequence 137, Application US/09907841
; Publication No. US20020198366A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same

;
;
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/907,841
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-907-841-137

Query Match 99.3%; Score 1986; DB 10; Length 382;
Best Local Similarity 97.9%; Pred. No. 2.7e-178;
Matches 374; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 MRPGTALQAVLLAVLLVGLRAATGRLLS-----GQVCRGGTQRPCYKVIYFHDTSR 52
Db 1 MRPGTALQAVLLAVLLVGLRAATGRLLSASDLRLGGQPVCRGGTQRPCYKVIYFHDTSR 60

QY 53 RLNFEEAKEACRRDGGQLVSI ESEDEQKLI EKFIENLLPSD GDFWIGLRREEKQSNSTA 112
Db 61 RLNFEEAKEACRRDGGQLVSI ESEDEQKLI EKFIENLLPSD GDFWIGLRREEKQSNSTA 120

QY 113 CQDLYAWTDGSI SQFRNMYVDEPSCGSEVCVVMYHQPSAPAGIGGPFYQWDDRCNMKN 172
Db 121 CQDLYAWTDGSI SQFRNMYVDEPSCGSEVCVVMYHQPSAPAGIGGPFYQWDDRCNMKN 180

QY 173 NFICKYSDEKPAVPSREAEGETELTTPVLPEETOEDAKKTFKESREAAALNLAYILIPS 232
Db 181 NFICKYSDEKPAVPSREAEGETELTTPVLPEETOEDAKKTFKESREAAALNLAYILIPS 240

QY 233 IPLL LLLVVTTVVCWVICRKRKREQPDSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSE 292
Db 241 IPLL LLLVVTTVVCWVICRKRKREQPDSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSE 300

QY 293 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYE 352
Db 301 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYE 360

QY 353 FSPDQMGRSKESGWVENEIYGY 374
Db 361 FSPDQMGRSKESGWVENEIYGY 382

RESULT 11

US-09-904-011-137
; Sequence 137, Application US/09904011
; Publication No. US20030003530A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc

```

; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/904,011
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 09/665,350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-904-011-137

```

Query Match 99.3%; Score 1986; DB 11; Length 382;
Best Local Similarity 97.9%; Pred. No. 2.7e-178;
Matches 374; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```

Qy 1 MRPGTALQAVLLAVLLVGLRAATGRLLS-----GQPVCRGCTQRPCYKVIYFHDTSR 52
Db 1 MRPGTALQAVLLAVLLVGLRAATGRLLSASDLRLGGQPVCRGCTQRPCYKVIYFHDTSR 60
Qy 53 RLNFEEAKEACRRDGGQLVSTSESEDEQKLEKFIENLLPSDGFWMGLRRREEKQSNSTA 112
Db 61 RLNFEEAKEACRRDGGQLVSTSESEDEQKLEKFIENLLPSDGFWMGLRRREEKQSNSTA 120
Qy 113 CODLYAWTDGSIQFRNMYVDPEPCGSEVCVVMYHQPSAPAGIGGYPYMFQWDDRCNMKN 172
Db 121 CODLYAWTDGSIQFRNMYVDPEPCGSEVCVVMYHQPSAPAGIGGYPYMFQWDDRCNMKN 180
Qy 173 NFICKYSDEKPAVPSREABGEETEELTPVLPEETQEEADAKTKFKESREAAALNLAYILIPS 232
Db 181 NFICKYSDEKPAVPSREABGEETEELTPVLPEETQEEADAKTKFKESREAAALNLAYILIPS 240
Qy 233 IPLLALLLVTTVVCWVWICRKRKREQDPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSE 292
Db 241 IPLLALLLVTTVVCWVWICRKRKREQDPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSE 300
Qy 293 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYE 352
Db 301 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYE 360
Qy 353 FSPDQMGRSKESGWVENEIYGY 374
Db 361 FSPDQMGRSKESGWVENEIYGY 382

```

RESULT 12

US-09-906-742-137

; Sequence 137, Application US/09906742

; Publication No. US20030023054A1

; GENERAL INFORMATION:

; APPLICANT: Genentech, Inc.

; APPLICANT: Ashkenazi, Avi

; APPLICANT: Botstein, David

; APPLICANT: Desnoyers, Luc

; APPLICANT: Eaton, Dan L.

; APPLICANT: Ferrara, Napoleone

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Fong, Sherman

; APPLICANT: Gerber, Hanspeter

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, A.

; APPLICANT: Godowski, Paul J.

; APPLICANT: Grimaldi, Christopher J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Hillan, Kenneth, J.

; APPLICANT: Kljavin, Ivar J.

; APPLICANT: Mather, Jennie P.

; APPLICANT: Pan, James

; APPLICANT: Paoni, Nicholas F.

; APPLICANT: Roy, Margaret Ann

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tumas, Daniel

; APPLICANT: Williams, P. Mickey

; APPLICANT: Wood, William, I.

; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

; TITLE OF INVENTION: Acids Encoding the Same

; FILE REFERENCE: 10466-14

; CURRENT APPLICATION NUMBER: US/09/906,742

; CURRENT FILING DATE: 2001-07-16

; PRIOR APPLICATION NUMBER: 09/665,350

; PRIOR FILING DATE: 2000-09-18

; PRIOR APPLICATION NUMBER: PCT/US00/04414

; PRIOR FILING DATE: 2000-02-22

; PRIOR APPLICATION NUMBER: US 60/143,048

; PRIOR FILING DATE: 1999-07-07

; PRIOR APPLICATION NUMBER: US 60/145,698

; PRIOR FILING DATE: 1999-07-26

; PRIOR APPLICATION NUMBER: US 60/146,222

; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-906-742-137

Query Match 99.3%; Score 1986; DB 11; Length 382;
Best Local Similarity 97.9%; Pred. No. 2.7e-178;
Matches 374; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 MRPGTALQAVLLAVLLVGLRAATGRLLS-----GQVCRGGTQPCYKVIYFHDTSR 52
Db 1 MRPGTALQAVLLAVLLVGLRAATGRLLSASDLRLGGQPVCRGGTQPCYKVIYFHDTSR 60

QY 53 RLNFEEAKEACRRDGGQLVSIIESEDEQKLIIEKFIENLLPSDGFWIGLRRRREEKQSNSTA 112
Db 61 RLNFEEAKEACRRDGGQLVSIIESEDEQKLIIEKFIENLLPSDGFWIGLRRRREEKQSNSTA 120

QY 113 CQDLYAWTDGSIQFRNWWYDEPSCGSEVCVVMYHQPSAPAGIGGYPMFQWDDRCNMKN 172
Db 121 CQDLYAWTDGSIQFRNWWYDEPSCGSEVCVVMYHQPSAPAGIGGYPMFQWDDRCNMKN 180

QY 173 NFICKYSDEKPAVPSREAEGETELTTPVLPEETQEDAKTKFKESREAAALNLAYILIPS 232
Db 181 NFICKYSDEKPAVPSREAEGETELTTPVLPEETQEDAKTKFKESREAAALNLAYILIPS 240

QY 233 IPLLLLVVTTVVCWVMICRKRREQPDSTKKQHTIWPSPHQGNSPDLEVNVIRKQSE 292
Db 241 IPLLLLVVTTVVCWVMICRKRREQPDSTKKQHTIWPSPHQGNSPDLEVNVIRKQSE 300

QY 293 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFTNDIYE 352
Db 301 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFTNDIYE 360

QY 353 FSPDMGRSKESGWENEIYGY 374
Db 361 FSPDMGRSKESGWENEIYGY 382

RESULT 13
US-09-906-838-137
; Sequence 137, Application US/09906838
; Publication No. US20030027143A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.

; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/906,838
; CURRENT FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 09/665,350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-906-838-137

Query Match 99.3%; Score 1986; DB 11; Length 382;

Best Local Similarity 97.9%; Pred. No. 2.7e-178;
Matches 374; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 MRPGTALQAVLLAVLLVGLRAATGRLLS-----GQPVCRGGTQRPCYKVIYFHDTSR 52
Db |||||
1 MRPGTALQAVLLAVLLVGLRAATGRLLSASDLRLGGQPVCRGGTQRPCYKVIYFHDTSR 60
QY 53 RLNFEEAKEACRRDGGQLVSIIESEDEQKLIIEKFIENLLPSDGDGFWIGLRRRREEKQSNSTA 112
Db |||||
61 RLNFEEAKEACRRDGGQLVSIIESEDEQKLIIEKFIENLLPSDGDGFWIGLRRRREEKQSNSTA 120
QY 113 CQDLYAWTDGSIQFRNWWYVDEPSCGSEVCVVMYHQPSAPAGIGGPFQWMDRCNMKN 172
Db |||||
121 CQDLYAWTDGSIQFRNWWYVDEPSCGSEVCVVMYHQPSAPAGIGGPFQWMDRCNMKN 180
QY 173 NFICKYSDEKPAVPSREAEGETELTTPVLPEETQEEADAKKTFKESREAAALNLAYILIPS 232
Db |||||
181 NFICKYSDEKPAVPSREAEGETELTTPVLPEETQEEADAKKTFKESREAAALNLAYILIPS 240
QY 233 IPLLALLLVTTVVCWWVICRKRKREQDPDPSTKKQHTIWPSPHQNSPDLEVNVIRKQSE 292
Db |||||
241 IPLLALLLVTTVVCWWVICRKRKREQDPDPSTKKQHTIWPSPHQNSPDLEVNVIRKQSE 300
QY 293 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYE 352
Db |||||
301 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYE 360
QY 353 FSPDQMGRSKESGWENEIYGY 374
Db |||||
361 FSPDQMGRSKESGWENEIYGY 382

RESULT 14

US-09-907-613-137
; Sequence 137, Application US/09907613
; Publication No. US20030027145A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/907,613
; PRIOR FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26

; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-907-613-137

Query Match 99.3%; Score 1986; DB 11; Length 382;
Best Local Similarity 97.9%; Pred. No. 2.7e-178;
Matches 374; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 MRPGTALQAVLLAVLLVGLRAATGRLLS-----GQPVCRGGTQRPCYKVIYFHDTSR 52
Db |||||
1 MRPGTALQAVLLAVLLVGLRAATGRLLSASDLRLGGQPVCRGGTQRPCYKVIYFHDTSR 60
QY 53 RLNFEEAKEACRRDGGQLVSIIESEDEQKLIIEKFIENLLPSDGDGFWIGLRRRREEKQSNSTA 112
Db |||||
61 RLNFEEAKEACRRDGGQLVSIIESEDEQKLIIEKFIENLLPSDGDGFWIGLRRRREEKQSNSTA 120
QY 113 CQDLYAWTDGSIQFRNWWYVDEPSCGSEVCVVMYHQPSAPAGIGGPFQWMDRCNMKN 172
Db |||||
121 CQDLYAWTDGSIQFRNWWYVDEPSCGSEVCVVMYHQPSAPAGIGGPFQWMDRCNMKN 180
QY 173 NFICKYSDEKPAVPSREAEGETELTTPVLPEETQEEADAKKTFKESREAAALNLAYILIPS 232
Db |||||
181 NFICKYSDEKPAVPSREAEGETELTTPVLPEETQEEADAKKTFKESREAAALNLAYILIPS 240
QY 233 IPLLALLLVTTVVCWWVICRKRKREQDPDPSTKKQHTIWPSPHQNSPDLEVNVIRKQSE 292
Db |||||
241 IPLLALLLVTTVVCWWVICRKRKREQDPDPSTKKQHTIWPSPHQNSPDLEVNVIRKQSE 300
QY 293 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYE 352
Db |||||
301 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYE 360
QY 353 FSPDQMGRSKESGWENEIYGY 374
Db |||||
361 FSPDQMGRSKESGWENEIYGY 382

RESULT 15

US-09-907-942-137
; Sequence 137, Application US/09907942
; Publication No. US20030027146A1
; GENERAL INFORMATION:

; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Shexman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/907,942
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 137
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-907-942-137

Query Match 99.3%; Score 1986; DB 11; Length 382;
Best Local Similarity 97.9%; Pred. No. 2.7e-178;

Matches 374; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
QY 1 MRPGTALQAVLLAVLLVGLRAATGRLLS-----GQPVCRGGGTQRPCYKVIYFHDTSR 52
Db 1 MRPGTALQAVLLAVLLVGLRAATGRLLSASDLDLRGGQPVCRGGGTQRPCYKVIYFHDTSR 60
QY 53 RLNFEAEACRRDGGQLVSIIESEDEOKLIEKFIENLLPSDGDWFGLRRREEKQSNSTA 112
Db 61 RLNFEAEACRRDGGQLVSIIESEDEOKLIEKFIENLLPSDGDWFGLRRREEKQSNSTA 120
QY 113 CQDLAWTDCSISQFRNWWYVDEPSCGSEVCVVMYHQPSAPAGIGGPPYMFQWDDRCNMKN 172
Db 121 CQDLAWTDCSISQFRNWWYVDEPSCGSEVCVVMYHQPSAPAGIGGPPYMFQWDDRCNMKN 180
QY 173 NFICKYSDEKPAVPSREAEGETELTPVLPEETOEDAKKTFKESREAAALNLAYILIPS 232
Db 181 NFICKYSDEKPAVPSREAEGETELTPVLPEETOEDAKKTFKESREAAALNLAYILIPS 240
QY 233 IPLLLLLVVTTCVWVICRKRREQDPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSE 292
Db 241 IPLLLLLVVTTCVWVICRKRREQDPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSE 300
QY 293 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYE 352
Db 301 ADLAETRPDLKNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYE 360
QY 353 FSPDQMGRSKESGWVENEIYGY 374
Db 361 FSPDQMGRSKESGWVENEIYGY 382

Search completed: December 22, 2003, 16:16:46
Job time : 36.7385 secs

CC AND THE CHONDROITIN SULFATE (CS) ATTACHMENT DOMAINS LIE BETWEEN G2
CC AND G3.
CC -!- PTM: CONTAINS MOSTLY CHONDROITIN SULFATE, BUT ALSO KERATAN SULFATE
CC CHAINS, N-LINKED AND O-LINKED OLIGOSACCHARIDES (BY SIMILARITY).
CC -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -!- SIMILARITY: Contains 4 link domains.
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC -!- SIMILARITY: Contains 1 Sushi (SCR) domain.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC -!- SIMILARITY: BELONGS TO THE AGGREGAN/VERSICAN PROTEOGLYCAN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U65989; AAB06238.2; -;
CC EMBL; S74662; AAC60527.1; -;
CC EMBL; L07054; -; NOT_ANNOTATED_CDS.
CC PIR; I46998; I46998.
CC HSSP; P08709; 1BF9.
CC InterPro; IPR000152; Asx_hydroxyl.
CC InterPro; IPR000742; EGF_2.
CC InterPro; IPR001881; EGF_Ca.
CC InterPro; IPR006209; EGF_like.
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR003006; Ig_MHC.
CC InterPro; IPR001304; Lectin_C.
CC InterPro; IPR000538; Link.
CC InterPro; IPR003324; SGXXSG.
CC InterPro; IPR000436; Sushi_SCR_CCP.
CC Pfam; PF00008; EGF; 1.
CC Pfam; PF00047; ig; 1.
CC Pfam; PF00059; lectin_c; 1.
CC Pfam; PF02339; SGXXSG; 66.
CC Pfam; PF00084; sushi; 1.
CC Pfam; PF00193; Xlink; 4.
CC PRINTS; PR01265; LINKMODULE.
CC PRINTS; PR00356; ANTIFREEZEII.
CC ProDom; PD000918; Link; 4.
CC SMART; SM00032; CCP; 1.
CC SMART; SM00034; CLECT; 1.
CC SMART; SM00179; EGF_CA; 1.
CC SMART; SM00445; LINK; 4.
CC PROSITE; PS00010; ASX_HYDROXYL; 1.
CC PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
CC PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
CC PROSITE; PS00022; EGF_1; 1.
CC PROSITE; PS01187; EGF_CA; 1.
CC PROSITE; PS50835; IG_LIKE; 1.
CC PROSITE; PS00290; IG_MHC; 1.
CC PROSITE; PS01241; LINK; 4.
KW Glycoprotein; Cartilage; Proteoglycan; Lectin; Signal; Sushi;
KW EGF-like domain; Repeat; Immunoglobulin domain.
FT SIGNAL 1 16
FT CHAIN 17 2333
FT DOMAIN 34 147
FT DOMAIN 170 247
FT DOMAIN 268 349
FT DOMAIN 513 590
FT DOMAIN 611 692
FT DOMAIN 2081 2117
FT DOMAIN 2130 2245
FT DOMAIN 2249 2307
FT DOMAIN 48 140
FT DOMAIN 152 247
FT DOMAIN 253 349
FT DOMAIN 495 589
FT DOMAIN 596 691
FT DOMAIN 694 816
KS.

FT	DOMAIN	819	1394	CS-1.
FT	DOMAIN	1395	2079	CS-2.
FT	DOMAIN	2080	2333	G3.
FT	DISULFID	51	133	BY SIMILARITY.
FT	DISULFID	175	246	BY SIMILARITY.
FT	DISULFID	199	220	BY SIMILARITY.
FT	DISULFID	273	348	BY SIMILARITY.
FT	DISULFID	297	318	BY SIMILARITY.
FT	DISULFID	518	589	BY SIMILARITY.
FT	DISULFID	542	563	BY SIMILARITY.
FT	DISULFID	616	691	BY SIMILARITY.
FT	DISULFID	640	661	BY SIMILARITY.
FT	DISULFID	2085	2096	BY SIMILARITY.
FT	DISULFID	2090	2105	BY SIMILARITY.
FT	DISULFID	2107	2116	BY SIMILARITY.
FT	DISULFID	2123	2134	BY SIMILARITY.
FT	DISULFID	2151	2243	BY SIMILARITY.
FT	DISULFID	2219	2235	BY SIMILARITY.
FT	DISULFID	2250	2293	BY SIMILARITY.
FT	DISULFID	2279	2306	BY SIMILARITY.
FT	CARBOHYD	126	126	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	239	239	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	333	333	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	387	387	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	444	444	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	620	620	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	676	676	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	747	747	N-LINKED (GLCNAC. .) (POTENTIAL).
SQ	SEQUENCE	2333 AA;	240573 MW;	8B9ED78F3508B596 CRC64;

Query Match 13.9%; Score 155.5; DB 1; Length 2333;
Best Local Similarity 28.5%; Pred. No. 1.4e-05;
Matches 51; Conservative 24; Mismatches 61; Indels 43; Gaps 11;

Qy	9	QPVCRCG---TQPCYKVIYFHDTSRRLNFEAEACRRDGGQLVSI ESEDEQKLI EKFI	65
Db	2120	QELCEEGWTKFQGH CYR--YFPD---RESWVDAESRCRAQOQSHLSSIVTPEEQ----	EFV 2170
Qy	66	ENLLPSDGD F-WIGLRRREEKQSNSTACQDLYAWTDGSI SQFRNHWYVDEPS---	CGSEVC 121
Db	2171	NN---NAQDYQWIGL-----NDR TIEGDFRWSGDGHSLOFENWRPNQDNFFVSGEDC	2219
Qy	122	VVM-YHQPSAPAGIGGYPYMFQWNDRCNMKNFICKYS----DEKPAVPSREAEGETE	175
Db	2220	VVMIWHEKG-----EWN DVP CNY YLPFTCKKGT VACGDP PVEHARTFGOKKD	2267

Search completed: December 22, 2003, 16:10:59
Job time : 6.56161 secs